UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II

Dottorato di Ricerca in Ingegneria Informatica ed Automatica

A FRAMEWORK FOR THREAT RECOGNITION IN PHYSICAL SECURITY INFORMATION MANAGEMENT

ALFIO PAPPALARDO

Tesi di Dottorato di Ricerca

XXV Ciclo

Aprile 2013

I Tutori

Prof. Valeria Vittorini

Dr. Concetta Pragliola

Il Coordinatore del Dottorato

Prof. Francesco Garofalo

Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione

Index of Sections

Index of Figures

Index of Tables

Introduction

In modern society, the capability to ensure an adequate level of security to persons and infrastructures is essential for the development of a territory. Malicious acts, including aggressions, intrusions, sabotages, and terrorist attacks as well as adverse natural events can pose a threat to the physical security. The protection against these threats is a need as well as a requirement in many application domains, including a wide range of industry and government sectors across the globe. Critical Infrastructure Protection (CIP) and Homeland Security (HS) are just a few of the possible examples. Whatever the application domain, the security of any asset – especially if it concerns complex, extended and critical environments – passes through the adequate use of protection technologies, techniques, tools, and methodologies, suitable for monitoring tasks and aimed at an intelligent surveillance.

The physical security of such environments requires the development of innovative approaches for identification, detection and mitigation of threats, vulnerabilities and risks. Therefore, it represents an area in which practical needs (e.g. coming from industry), technological resources (e.g. belonging to physical security market) and scientific research (e.g. on information fusion strategies and event correlation techniques) converge all together .

The events of September 11, 2001 brought a rapid expansion of research efforts in that direction, in particular to prevent terrorist acts, minimize the damage and recover from disruptive events. Infrastructure protection against potential threats is usually performed by surveillance systems that are more and more large, distributed and heterogeneous. The cyber-physical and human-in-the-loop nature of this field requires a set of multidisciplinary activities to be performed in order to adopt appropriate and effective protection mechanisms. Due to the variety of natural and malicious threat scenarios, a growing set of different sensing technologies is required. However, many of the developed innovative technologies (e.g. video analytics) do not always provide adequate intelligence and reliability. On the contrary, the evolution of end-user requirements increasingly calls for enhanced early warning capabilities and superior situation awareness, which the traditional technologies and systems cannot provide.

An effective information integration and management is essential to overcome technological limitations, synthesizing data from multiple alerting systems and physical sensors. Thanks to appropriate methodologies and techniques, the exploitation of distributed and heterogeneous sensorial subsystems (encouraged also by the development of novel low-cost devices) can lead to several levels of event correlation. At low level, the development of multi-modal approaches for monitoring and surveillance activities helps in providing advanced event detection capabilities and/or in improving detection reliability. At a higher level, information aggregation is also the key to develop the next generation of security management systems, the so called PSIM (Physical Security Information Management) systems. PSIM systems help to integrate security devices, to improve detection efficiency and effectiveness, and to produce an increased situation awareness. The main factor to achieve those results is the presentation of all the relevant information into a single view together with essential decision support features.

At the same time, the protection of many infrastructures – which could be “open”, spread through hundreds of kilometres, and vulnerable to many threats of various kind and seriousness – may require hundreds or thousands of cameras and other sensors, which makes human-based surveillance unfeasible. Furthermore, the detection of specific events or activities almost completely relies on costly and scarce human resources. Manual analysis of video as well as diverse sensor alarms (which can be false) is labour intensive, fatiguing, and prone to errors. Additionally, psychophysical research indicates that there are severe limitations in the ability of humans to monitor simultaneous signals. Thus, it is clear that there is a fundamental contradiction between the current surveillance model and human surveillance capabilities.

In such a context, the novel research must aim at recognizing threats scenarios as early as possible, providing superior situation awareness and decision support, in order to quickly – possibly automatically – activate response-and-recovery strategies. The thesis addresses that issue on different levels:

- At a methodological level, we have proposed the proper use of ad-hoc information fusion strategies, which contribute to define the general and challenging paradigm of “augmented surveillance”. Its declination in a specific domain should help in exploiting technologies, capabilities and tools (already available in the state-of-the-art), in order to perform functions of different complexity.
- At the application level, we have developed a framework aimed at the automatic and early detection of threats against infrastructures, by performing a model-based logical, spatial and temporal correlation of basic events detected by the sensorial subsystems. The design of a proper detection engine is the key to set up an effective reasoning on heterogeneous data and to implement an application of fusion. It is also the result of the search for a light, efficient and easy-to-use approach, obtained by properly selecting models and correlation techniques.

This thesis is structured as follows. Chapter 1 provides the state-of-the-art and open issues in information integration, fusion and management in the specific context of physical security. The chapter describe the main resources available, coming from the industry and scientific research fields, to address the posed problems.

Chapter 2 describes the main requirements of the methodological approach to the physical security. The fulfilment of such requirements is fundamental to conceive solutions that should be effective as well as viable. The chapter also describes how to use the information fusion strategies in defining the “augmented surveillance” paradigm.

Chapter 3 introduces the basic motivation, the working logic and the architecture of DETECT (DEcision Triggering Event Composer & Tracker), the framework we have developed for the automatic detection of the physical security threats, possibly before they can evolve. It operates by performing a model-based logical, spatial and temporal correlation of basic events, detected by each monitoring device or subsystem.

Chapter 4 and Chapter 5 address the main limitations of the approach on which DETECT is based. In fact, it is deterministic and takes into account a pre-defined knowledge base. Therefore, in order to improve the detection effectiveness, a heuristic recognition, based on the computation of "distances" between threat scenarios, has been proposed. The distances are defined using ad-hoc metrics for the detection models of DETECT. At the same time, in order to improve the detection effectiveness, we have described how to handle and elaborate the main sources of uncertainty, which include sensors and models. The aim is to quantify the detection reliability level, in order to improve the decision support in triggering response actions and to control the rate of false alarms.

Chapter 6 presents some applications of the additional features provided by DETECT in a specific field, like the railway and mass-transit domain. More in detail, it includes examples of threat modeling and real-time computation of distances and alarm reliability for some threats of reference.

Chapter 7, finally, describes the working principles and advantages of an overall integration of DETECT with a PSIM system. In particular it presents the integration with the existing Security Management System (SMS), developed by Ansaldo STS. It represents the concrete attempt to solve a well-known class of problems, which involve most of the actors moving in the physical security landscape.

Chapter 1

State-of-the-art and open issues in information integration, fusion and management for physical security

Security is "*the state of being free from danger or threat*". According to that definition provided by the Oxford English Dictionary, the security is a condition related to the degree of protection from threats, concerning any person, community, organization, or nation. More in detail, the concept can be applied to any asset, which has a certain vulnerability to given threats and therefore a certain level of risk. In modern society, the development of a territory is increasingly tied to the capability to ensure an adequate level of security to persons and infrastructures. Criminal acts, including aggressions, sabotages, and terrorist attacks as well as accidents and adverse natural events can pose a threat to the physical security. The protection against these threats is a need as well as a requirement in many application domains.

Today, more than ever, the welfare, the quality of life and all the vital functions of a country increasingly depend on the continuous and coordinated operation of several infrastructures, which for their importance are defined as Critical Infrastructures (CIs). The main CIs belong to Transport, Energy and Telecommunication networks, however the list include a disparate set of sectors, depending on the indications of each government (in the European Union area, the guidelines for all the member states are defined by the European Commission). The physical security of such infrastructures requires the development of innovative approaches for identification, detection and mitigation of threats, vulnerabilities and risks. The same approach involves a wide range of industry and government sectors across the globe: Homeland Security, Defense, Law Enforcement, Corporate Security, etc. The events of September 11, 2001 brought a rapid expansion of Critical Infrastructure Protection (CIP) efforts, in particular to prevent terrorist acts, minimize the damage and recover from disruptive events.

Whatever the application domain, the security of any asset – especially if it concerns complex, extended and critical environments – passes through the adequate use of protection technologies, techniques and methodologies, suitable for monitoring tasks aimed at an intelligent surveillance. The description which follows is biased in that direction.

1.1 Integration and management of physical security systems

The evolving requirements to be fulfilled in surveillance applications increasingly call for the deployment of intelligent and automated monitoring solutions. Also, the growing capabilities assured by the technological progress encourage the employment of advanced surveillance systems. Although nowadays video surveillance represents the most popular form of surveillance, there are many other forms of monitoring. Given the size and complexity of the sensed environments, it is easy to understand that modern surveillance cannot be performed via cameras only, but should include multiple modalities. Each monitored area offers different information streams, which need to be captured, evaluated and possibly correlated. According to the specific application, an ad-hoc set of heterogeneous sensing technologies is required.

Today the physical security industry is complex and reasonably mature. The main technologies include video surveillance (possibly including video analytics), intrusion detection, access control, intelligent sound detection, CBRNe (Chemical Biological Radiological Nuclear explosive) agent detection, and many others [30]. On the other hand, most basic systems can include for example temperature, humidity or pressure sensors. However, with respect to the information integration and management, the industry is still underdeveloped. From this point of view, the scientific community is recognized as a valuable resource for the introduction of new solutions.

In practical applications, each monitoring system is handled by means of an ad-hoc software platform. The traditional surveillance solutions include, for example, VMS (Video Management System), ACS (Access Control System), etc. They provide an overview of the installed devices (with a related report of diagnostic, warning, and alert messages) and a set of basic functionalities (e.g. for data acquisition, control, configuration, rules setting). However, in this way each event is treated separately with a lack of an effective information sharing. Therefore, the result is a very fragmented approach to the physical security. Assuming the use of multiple security systems and a remote surveillance of each of them from a control centre, the possible consequences are the following:

a. a human operator at the control centre may be inundated of warning messages, coming from multiple separated interfaces (one for each management system of the single technology);
b. given the limited detection reliability of each triggered alarm (due to the limited reliability of any technology), it is necessary as well as complex to evaluate and confirm each of them;
c. for the ones confirmed, the capability in quickly understanding the criticality level (and to know if there is one more critical than others) is very low;
d. for each detected alarm, the quick activation of ad-hoc response procedures (possibly considering first the ones with higher priority) is not adequately supported.

As one can see, within the situation management task, the events contextualization is crucial. Generally speaking, video streams can provide a context (and then a "meaning") for all security data. For example, in a public building, if the intrusion detection system detects an event in correspondence of an emergency exit, it may be associated to a person who exited the door unintentionally or to a terrorist who is preparing an attack. In fact, the system provides an alert, but no context. Only viewing the video stream related to the camera at the emergency exit, the severity of the situation can be evaluated and a countermeasure can be adopted. Although such a principle is intuitive, its practical application in traditional surveillance approach is not so immediate.

1.1.1 Physical security information integration

Current research tends to combine and to exploit multiple modalities of monitoring, in order to create distributed surveillance solutions, including not only multi-camera systems [3], but more in general multi-sensor surveillance systems [83]. The combination of traditional video surveillance with other smart sensing technologies (see also [72]) leads to the development of new multimedia surveillance systems [20], which collect and process different information streams (audio, video, and any output of a sensor). To that aim, a proper integration (e.g. by means of diverse algorithms) of these information is required [30]. Therefore, a multimodal and multimedia solutions combine two characteristics:

- the use of multiple sensors, possibly with overlapping sensing areas and which communicate between them through a network;
- the use of heterogeneous sensors, to exploit all the information available in the monitored area.

The most advanced forms of multimedia and multimodal surveillance are the answers to different needs: to overcome specific limitations of the single modality (e.g. video analytics algorithms suffer from several problems); to improve the event detection reliability, which is crucial in determining viability and effectiveness of surveillance systems; to extend the capabilities in detecting complex events, considering also the evolution of end-user requirements.

The integration of heterogeneous sensors within a multimedia and multimodal surveillance system is the key to address the issues described above. Regarding the integration approaches, they can be classified in two categories: *bottom-up* and *top-down*. Bottom-up approaches are aimed at developing ad-hoc integration algorithms whose inputs are the outputs of the sensors to be integrated. In such a case, the algorithm executed by sensors is written without using information coming from other sensors. Hence the whole integration logic lies at a higher level of abstraction. Typically, the approach addresses the need for improving the overall performance of the surveillance system by trying to reduce false alarm rate (named FAR), and to add functionalities. Top-down approaches are aimed at developing the algorithms executed by the sensors, using the information coming from other devices. In this case, sensor output depends on the presence of other sensors and on the information they provide. Therefore, there is a part of the integration logic to be implemented in each sensing device, i.e. at a lower level of abstraction. Typically, this approach is conceived to improve reliability of detection (represented by a parameter named POD, Probability Of Detection). For example, in video analytics applications, the object detection and tracking performed by the single camera can be improved by means of additional information from other cameras or sensors.

Some examples of multimedia and multimodal surveillance are described as follows. Audio surveillance (see also [62], [63], and [64]) can be auxiliary to video surveillance in solving specific problems like the tracking of people in case of occlusions (correlation among unobserved audio and video) or the identification of a region of interest by a camera (correlation among observed audio and video) [48][77].

A network of PIR (Passive InfraRed) sensors can support object and motion detection, performed by the video analytics. This is particularly important when the classic methods based on shape and colour recognition fail because of the limited field of view of cameras, or when they are deployed in places with very different lighting conditions. Thus, PIR sensors can be used to detect motion with a high accuracy (also during the night), since they are not sensitive to light conditions [66].

Furthermore, laser technologies like the LDV (Laser Doppler Vibrometer), can be used to remotely capture acoustic signals like human speech (since performs a vibration detection within two hundred meters). An integrated system with multiple cameras and LDVs can represent an advanced form of multimodal surveillance for face and voice recognition systems [67] [84]. Another laser based system that comes in support to the video analysis is the LIDS (Laser Intrusion Detection System). For example, it can be applied to protect areas from intrusions, and to monitor of portals and tunnels in order to detect unauthorized objects or people. By managing known profiles, which can be recognized by the system without triggering a warning, LIDS is able to reduce the rate of nuisance alarms. The insensitivity to lighting conditions, reflective surfaces, rain, and snow has a great impact on reducing the false alarm rate with respect to video surveillance. Possible applications include also platform line crossing and intrusions (accidental or malicious) onto the track in a railway context. Therefore, these systems offer both an effective support for visual analysis and additional information which can be correlated with data coming from intelligent video surveillance [41]. A survey on multi-sensor integration for wide-area surveillance is provided in [1].

As described, multimodal monitoring forms represent suitable solutions for complex and/or crowded environments. However, they don't represent the cure for everything. Besides, some benefits are counteracted by problems, like the harder camera calibration and configuration (to adequately fulfil the overall task); and the more complex management of available cameras, which is fundamental in performing functions like the object tracking from one camera to the next, when it is necessary to establish a correspondence using common reference points.

In many practical applications, the multiple sensing systems are integrated at junction-box level, the first point of aggregation of the detected signals. The result is a low level integration, which allows a more immediate and reliable data correlation, but that is characterized by several disadvantages. The integration may be more difficult for the different representations (and semantic levels) of the combined modalities (e.g. audio and video), while the scalability of the solution is very limited.

1.1.2 Physical security information management

The limitations related to the traditional management systems of the single technologies (e.g. VMS, ACS, etc.) and to the multimodal monitoring solutions, highlight the need for a different – and possibly more cohesive – approach to address the overall information integration and management, especially in the context in which we move.

Regarding the integration, if each system works within their own confines to perform a defined task and there is no knowledge of each other, the absence of interactivity leads to a limited effectiveness in the use of available information. On the other hand, with a limited capability to manage technologies and related information flows, as end-users continue to add data into the security information flow, the more they have reduced means to manage and to use that data. Thus, the flow is increasingly big and heterogeneous, while the security departments (regardless of public or private organization) get larger and more inefficient. The consequences are a higher cost for the security, which is increasingly more difficult to justify, and an organizational inefficiency, which leads to a false sense of security.

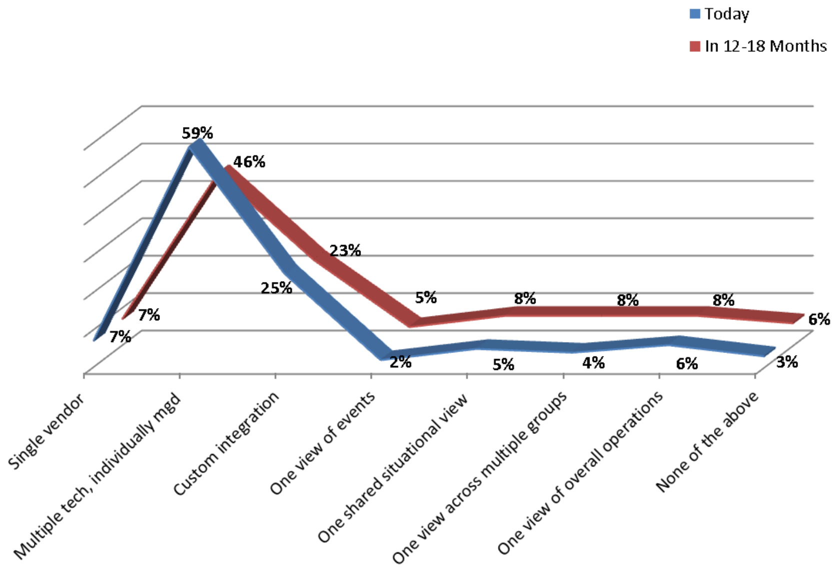


Figure 1 – Progression path in physical security information management

A real picture of current situation is in Figure 1, which shows the progression path of security and technology maturity through seven stages, beginning with a single vendor environment and evolving to a system that provides a complete view of situations (also across multiple organizations). As reported in figure, in most of the cases (almost 60%), public and private organizations are still in the second stage, i.e. the physical security is addressed using multiple technologies, individually managed [79]. In such a scenario, to get the right information to the right people at the right time is not possible.

Ad-hoc information aggregation and management are paving the way to a new generation of PSIM (Physical Security Information Management) systems, capable to address the following requirements [68]:

- device level information from a wide set of disparate security systems, which may incorporate products of more independent manufacturers, should be gathered;

- given the possible size and complexity of infrastructures, a great amount of information should be evaluated in order to recognize and prioritize critical situations;
- the overall monitored situation should be presented to human operators in a clear, concise and comprehensive format, enabling an accurate and quick confirmation of the alarms;
- procedures and step-by-step instructions should be triggered to respond to the confirmed alarms;
- all the activities should be tracked and recorded to aid compliance management and to enable post-event investigative analysis.

They represent the five key functionalities that a PSIM system should include, in order to provide a complete Situation Assessment and Situation Management, to effectively manage any security-related event or emergency in real-time and across any organization.

After the terrorist attack on the 11th September 2001, it was evident that the correct interpretation and the consistent response, with respect to many critical events at the same time, could not be left to human operators only. In particular, the emergency response of the traditional security management systems had to be improved. The PSIM acronym was born approximately in 2005, to tag the systems able to support the analysis and the automated decision making, and not only to gather and present information coming from the devices, like in the traditional approaches [68]. Obviously, the quick and punctual activation of response procedures, depending on priorities and criticality levels of the events, can significantly improve the overall impact of countermeasures.

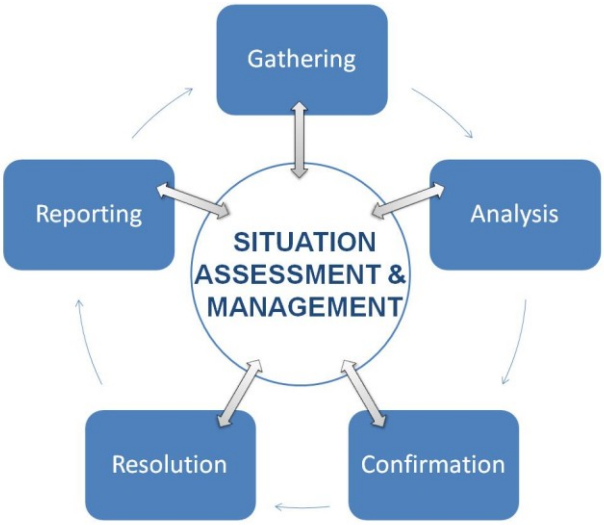


Figure 2 – PSIM key capabilities

As already mentioned, the main capabilities of a PSIM system include: a) gathering of data and information from the field; b) analysis and interpretation of received data, events and alarms; c) confirmation about the authenticity of the alarms; d) resolution of critical situations and possible emergencies in real-time; e) reporting of all the tracked information (see Figure 2).

The basic concept behind PSIM is not new, in fact it applies the experience from the software and network security areas to the physical security to optimize devices integration, analysis and end-to-end situation management and resolution. More in detail, PSIM is analogous to SIEM (Security Information and Event Management) software. It does for physical security what SIEM does for cyber security, simplifying the surveillance task, while improving security and reducing the time, cost and effort that physical security requires [80].

High level architecture

One of the key objectives of a PSIM system is to integrate diverse systems into a common information model. From this viewpoint, PSIM provides a platform to connect any number and type of security devices or systems and advanced processing capabilities of the device information. A plug-and-play approach with robust SDKs (Software Development Kits) and APIs (Application Programming Interfaces), make the integration with new subsystems (e.g. intelligent video surveillance, biometric access control, CBRNe detection, etc.) quick and seamless.

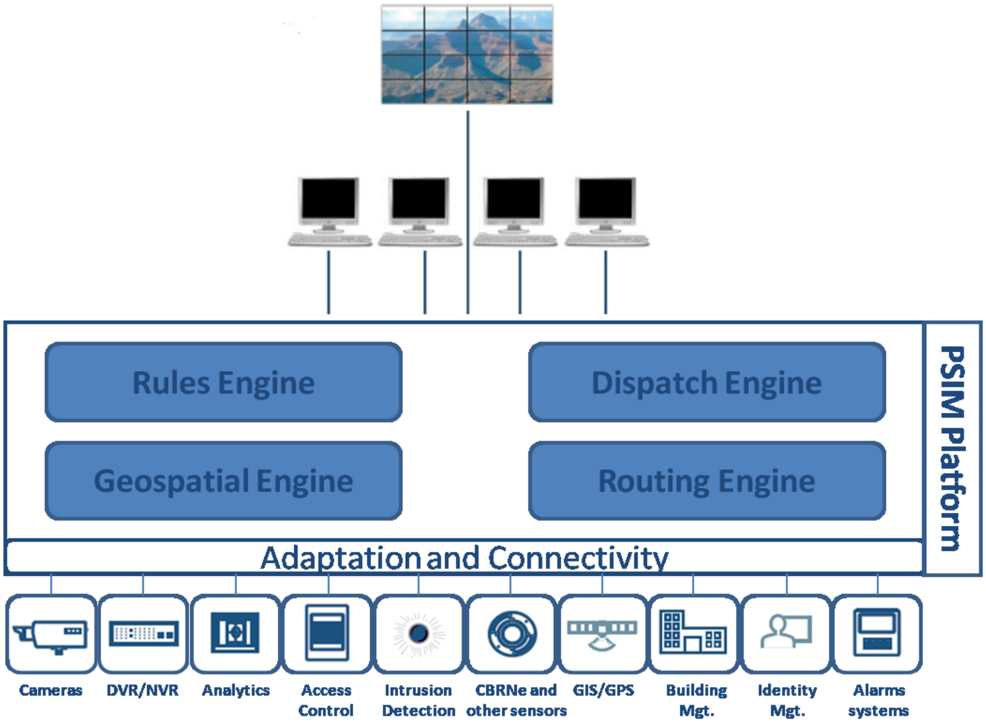


Figure 3 – Modular PSIM architecture

In addition to that, the key components of a PSIM solution are the engines that translate data (from physical security subsystems) into intelligence [53]. These engines represent the main resource to identify, analyze and prioritize situations. They also ensures that all appropriate information is presented to an operator in an integrated display, aiding him/her in focusing on the situation (e.g. the possible threat) and not on the technology. From this point of view, with reference to one of the most common technologies, it is worth noting that PSIM incorporates, but also transcends the video alerts.

Figure 3 shows a generic architecture of a PSIM platform, which enable the connection with a disparate set of security devices through pluggable modules [80]. They can range from video surveillance components, access control, CBRNe sensors, to other systems that are in a certain way security-related, although they are not strictly security systems. Some examples are: GPS (Global Positioning System), Fire alarm, Building Management (e.g. elevator, HVAC[1], lighting), etc. The Adaption and Connectivity layer receives data through various protocols, SDKs, and APIs and creates a "common language" for security information. Such a language enables the processing performed at higher levels. Obviously the type of adaptation depends on the type of outputs of the integrated sensing subsystems, and on their possible capability to perform this task on their own.

The translation of data into intelligence is possible thanks to ad-hoc processing and correlation of captured information. According to the specific application, a set of engines can support the task. For example, a geospatial engine can provide location awareness of devices and support mapping functionalities (e.g. rules can be set such that multiple alarms from one location can be correlated). A routing engine can optimize the use of the network resources. A rules engine can analyze and correlate events from multiple sources, in order to infer new knowledge about the overall situation and better support decisions. A dispatch engine can activate external transmission of messages and commands, also depending on the indications of rules engine, since it executes recommendations for situation resolution. Finally, information integration and management imply a number of activities which are usually performed by means of a set of tools, e.g. including display and video wall systems. All the equipments can be effectively managed by a unified user interface, thereby making situation assessment and management simpler, and reducing reaction time in case of critical events.

[1] The acronym HVAC stands for Heating, Ventilation and Air Conditioning

[2] An example of guideline is provided by the chart "PSIM vs. VMS: what do you need?" prepared by

“True” and “Lite” PSIM

PSIM represents a form of integrated management, but not all the forms of integrated management correspond to a PSIM solution, in particular because the final objective is to integrate many independent subsystems (possibly of different manufacturers) which were not designed to be integrated. The management systems of the single technologies (e.g. VMS, ACS, etc.) have developed quickly in recent years and they represent very powerful security solutions. However, it is important to explicit the difference between these “lite” solutions and the “true” PSIM ones, which feature (at least) the five key capabilities in Figure 2. One of the main doubts about PSIM’s adoption is precisely due to the confusion on what it is and what it does [46]. Part of the reason is that VMS or ACS vendors tend to blur the confines between “lite” and “true” PSIM, in the attempt to extend the market of their own products [68].

Many VMSs have interfaces to other physical security systems (e.g. access control, intrusion detection, fire alarms), but most of them don’t seamlessly integrate with competing software. The “true” PSIM solution is an open system, independent from the specific vendors of the technologies. If a PSIM system was totally based on an certain VMS/ACS product able to integrate other systems, its degree of openness would be highly reduced. In fact, it is quite obvious that a system manufacturer will not share its know-how with a company which could be its competitor. For example, if the PSIM vendor also produces cameras, it is quite difficult to find a further third party provider for cameras and/or VMS to be integrated. Therefore a “lite” PSIM system (also named Tier 2 or 3 systems [36][68]) could became isolated over the years, and the end-users could not have the interoperability they believed to have. On the other hand, a “true” PSIM system should integrate most of physical security systems. The interfaces are typically “one-and-done”, i.e. once developed they are part of a library. A critical issue can arise when the provider has to develop them for the first time, because of the related cost. At the same time, end-users don’t like to be the first ones for which new interfaces are deployed (e.g. they prefer more consolidated integrated platforms). The consequence is that PSIM vendors tend to develop their market in specific verticals: transportation (railway, airport and seaport), defense, homeland security or other critical infrastructures. Therefore, the main differences between the vendors are in their expertise in a specific application domain and in supporting specific security systems for that domain.

In addition to that, the lack of the lite solutions in data analysis complicates the decision making of human operators, and their reactivity. In such a process, speed and accuracy are crucial to avoid service interruptions, dangerous domino effects, financial damages, and to save assets or even lives.

Finally, unlike lite systems, PSIM systems have a high scalability. In extended infrastructures, this means also a greater capability in providing an effective resilience, service continuity and crisis management.

Despite the added value that a PSIM system should have, the mentioned aspects don't mean that a "true" PSIM solution is the best in any case. Of course it depends on the requirements coming from the specific end-user, which operates in a certain application domain. For example, there are ad-hoc guidelines (for end-users) to determine whether a VMS or a PSIM is a better fit according to the needs[2]. From the functional point of view, the following points allow for a proper comparison:

- PSIM may depend on video streams provided by a VMS, but it doesn't substitute for a VMS. PSIM doesn't record or manage video, but can manage situations and possibly critical events;
- VMS can offer a basic form of situation awareness, e.g. if it is integrated with an ACS. In that case VMS is able to report an alarm from a camera or an access control device. Once received it is acknowledged by an operator and the process stops. From this point, unlike VMS, PSIM begins to support a real-time situation awareness and management as the criticality evolves (e.g. it can

[2] An example of guideline is provided by the chart "PSIM vs. VMS: what do you need?" prepared by Bob Banerjee, NICE Systems

correlate the incoming alarms with the previous ones, instead of issuing a new notification, and automatically associate an ad-hoc emergency procedure to resolve the criticality);

- PSIM can address compliance issues (e.g. with respect to CIP programs and regulations), using automated workflows and incident reporting capabilities. At the same time, unlike VMS/ACS, it can ensure business continuity and help in implementing contingency plans;
- PSIM could not be exclusively dedicated to security-related events. In fact, more in general, it can enable an operational situation awareness.

1.2 Information fusion strategies

1.2.1 The need for an overall process model

Multimodal and multimedia surveillance is aimed at providing complementary information and at increasing the accuracy in detecting threats and/or events of interest. Indeed, it is easy to understand that by exploiting multiple features from the monitored area, the power of event detection is greater than the one assured by a single source. However, in order to recognize in real-time complex situation patterns, to build hypotheses of unfolding situations, and to take actions in response to these situations, the overall capabilities of the surveillance system should include processing, correlation and handling of multimedia data coming from different sources.

At the same time, because of the variety of natural and malicious threat scenarios, a growing set of different sensing technologies can be required. Unfortunately, many of the recently developed innovative technologies (e.g. video analytics) do not always provide adequate reliability (see e.g. [40][55]). Many automatic detection systems generates unnecessary warnings, which can be classified as false alarms or nuisance alarms. Especially with regard to the decision support feature of surveillance systems (e.g. for triggering countermeasures), it is very important to control the rate of these alarms (see e.g. [13]).

The integration of information coming from different sources represents the basic concept behind the new generations of surveillance solutions, where many different media streams contribute to provide a greater situational awareness, an improved early warning capability, and a better decision support. Whereas the capabilities of the traditional systems are limited in data analysis and interpretation, and hence in real-time prevention and reaction. Furthermore, since a few human operators are usually employed in security surveillance, human-factors also need to be carefully addressed, including cognitive ergonomics in human-machine interaction [9][82].

Regardless the specific system to be implemented, the first objective is to model the overall integration process of heterogeneous information, in order to conceive a real-time data comprehension framework. Within the context of analysis and reasoning about dynamic situations, where application domains include information-rich environments, an active field of research is focused on Information Fusion (IF). IF takes into account the specific aims related to the application domain, on the one hand, and the different characteristics of the available multimodal subsystems, on the other. In fact, IF may: provide information at different semantic levels and in different formats, require different kinds of processing, have different reliability levels, and have a certain degree of mutual correlation [6]. The complexity of such a task requires an appropriate strategy to fuse the available information. Using an efficient fusion scheme, one may expect significant advantages, such as:

- Enlarging information extraction from the available sources;
- Improving confidence in decisions by leveraging more information;
- Increasing robustness against sensor failures and outliers in measurements (stability).

However, an important issue regarding IF is that, while using additional information is intuitively advantageous to add knowledge and to support decisions, the overall performance of the fusion process can decrease in case of additional incorrect data [23]. Other basic concepts characterizing IF systems and the models proposed in literature are described in the following.

1.2.2 Information Fusion models

Data need to be analyzed effectively and efficiently to provide appropriate information for intelligent decision making [78]. Hence, the power of Information Fusion is being increasingly considered in several applications. Empirical studies have shown the overall improvements of information systems based on fusion of different information sources [42]. In particular, fusion of relevant data has proven effective in reducing uncertainty (e.g. false alarm rates), in increasing accuracy (in terms of confidence levels) in the early detection of threats, and in increasing robustness by exploiting redundant information [14]; being able to deal with data that is redundant, inconsistent and conflicting [4] is also essential. The basic motivation of IF is described as follows: *"exploiting the synergy in the information acquired from multiple sources (sensor, databases, information gathered by humans, etc.) such that the resulting decision or action is in some sense better [...] than would be possible if any of these sources were used individually without synergy exploitation"* [24].

Although it is widely recognized that IF can support and enhance decision making, an Information Fusion System (IFS) is not concretely viewed as a Decision Support System (DSS) [60]. In this sense, many heterogeneous fields of research often exploit the results already available in other sectors like defense [81].

Several works have attempted to characterise IFS, but actually there is no general consensus in the literature regarding the components of an IFS; consequently, there are slightly different opinions on what is required for a system to be classified as an IFS.

Basically, we can say that an IFS needs to receive information from different information sources, including sensors and smart devices, human sources or data archives (depending on the context). The sources could be classified as either past (e.g. data archives), present (e.g. sensors) or future (e.g. simulations/models) [60].

An information fusion process of different sources can be automated with the purpose to achieve timely, robust, and relevant assessment of unfolding situations (e.g. in terms of threats, within the context of physical security) and their possible projections.

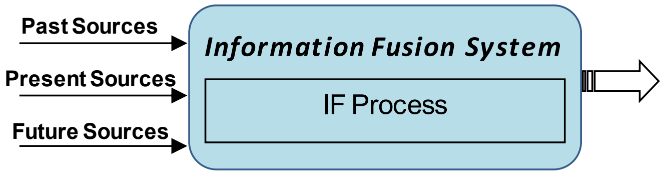


Figure 4 – Schematic representation of the information fusion system

Recently, it has been acknowledged that the user could actually contribute to the information fusion process [11][12][59]. Typically an IFS involves different degrees of automation and user involvement within two extremes: ‘user dominant’ (i.e. user is in control of the fusion process) and ‘machine dominant’ (i.e. fully automated fusion process).

In the IF research community, different models have been proposed to have a common understanding across different applications domains which use the information fusion concepts. The most significant ones are presented in the following.

The JDL model

The following model was created by the U.S. Joint Directors of Laboratories, hence the name ‘JDL’. It is the most commonly used model which categorises the fusion process. In general, the model describes how IFS transforms sensor data into information which a user can employ for decision making [57].

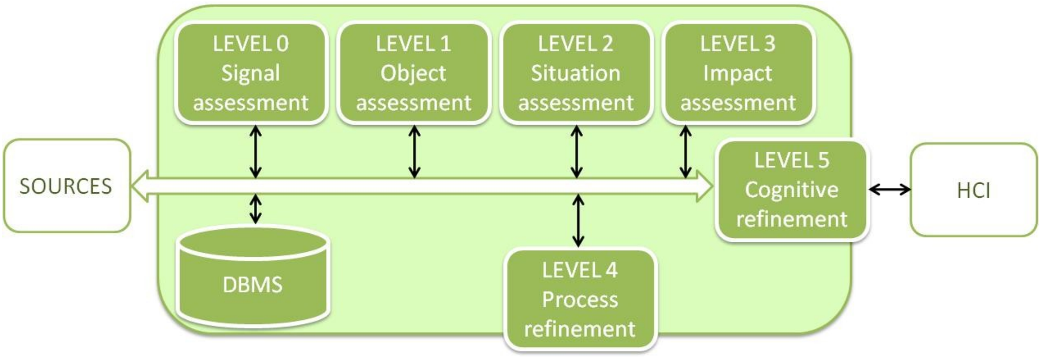


Figure 5 – JDL model

The model is readable from left to right (see Figure 5), from the different sources of information to the user interface, i.e. the HCI (Human Computer Interface). Between them, the different levels may be viewed in a hierarchical order, although the JDL model is not a process model indicating a flow. Rather it shows different categories of functions. While the DBMS (DataBase Management System) supports the maintenance of the data used and provided by the IFS.

Only for convenience, the functions are described in their hierarchical order. In particular, levels 0-3 represent the "assessment functions", instead of levels 4 and 5 which represent "refinement functions". The latter could be considered as a sort of meta-processes, which control and refine the previous levels. More in detail, the different levels are described as follows [42]:

- Level 0 pre-processing (*signal assessment*): this level pre-processes data at the individual sensor in order not to overwhelm the system with raw data;
- Level 1 processing (*object assessment*): "fusion of multi-sensor data to determine the position, velocity, attributes, characteristics, and identity of an entity (such as an emitter or target)";
- Level 2 processing (*situation assessment*): "automated reasoning to refine our estimate of a situation (including determining the relationships among observed entities, relationships between entities and the environment, and general interpretation of the meaning of the observed entities)";

- Level 3 processing (*impact assessment*): "projection of the current situation into the future or define alternative hypotheses regarding possible threats or future conditions". This level is also sometimes referred to as threat refinement/assessment;
- Level 4 processing (*process refinements*): "a meta process that monitors the ongoing data fusion process to improve the processing results (namely improved accuracy of estimated identity of entities and improved assessment of the current situation and hypothesized threats)";
- Level 5 processing (*cognitive refinements*): "interaction between the data fusion system and a human decision maker to improve the interpretation of results and the decision-making process".

First of all, raw data may be pre-processed (Level 0-signal assessment) in order to assess the signals from the sensors and extract key information (e.g. functions such as video, audio, or signal processing). Since the information sources could refer to sensors as well as agents (human sources) or data archives, this activity should be tailored on the typology of the sources, bringing the extracted information to the same semantic level before the subsequent processing.

The second function is object assessment (Level 1) and it concerns the combination of data from different sources to obtain estimates of an object's attributes or identity (e.g. classical techniques such as tracking and pattern recognition are used).

Level 2, situation assessment, is a collection of functions to interpret the different objects' relationships and their relationships with the environment (typically automated reasoning and artificial intelligence techniques are used here). The difference between the two levels is the following: Level 1 involves attribute-based state estimations, while Level 2 involves relation-based state estimations [44].

Impact assessment (Level 3) concerns the future states and the projection of the interpreted situations, in order to assess the possible threats, risks and impacts.

In Figure 5, Level 4 and Level 5 are located on the border of the fusion process. They are quite similar, although there are some distinguishing features. The main difference between them lays in the responsibility of the refinement process: in Level 4, the responsible is the system itself, in Level 5 it is the user who controls the process depending on the particular needs he/she has at the moment. Anyway, the incorporation of Level 5 into the JDL model has not yet achieved common usage within the information fusion community [43]. However, the aspect related to the understanding of the active role of human information processing in IF should be carefully addressed [58]. The JDL model is under constant revision, and although other models have emerged, they have not gained the same popularity. One of the reasons for that is related to the holistic perspective provided by the model, which is usable for many purposes related to the research domain of IF systems [61].

Dasarathy's functional model

Dasarathy defined a useful category of different fusion functions, based on the types of data and information processed and on the types of results obtained from the process [22]. The input and output of a fusion process can be of any level: Data, Feature, and Decision. For this reason the Dasarathy's functional model is also known as DFD model (see Figure 6).

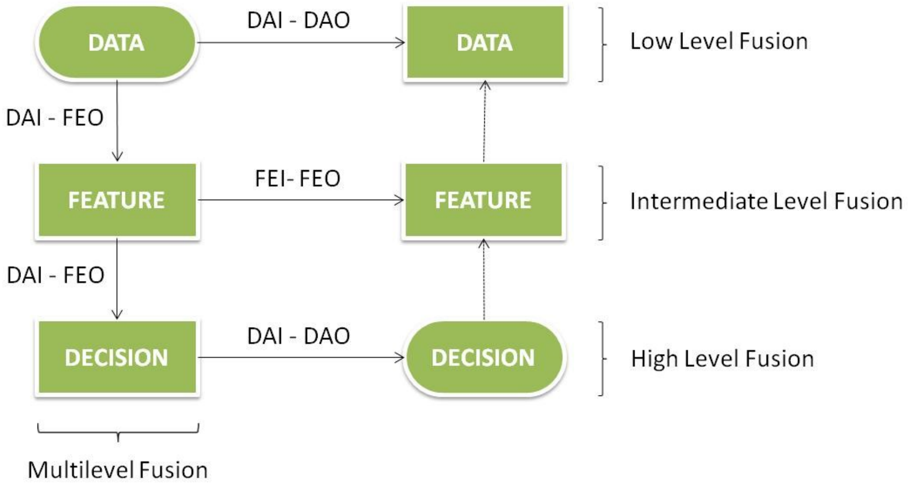


Figure 6 – DFD model

In this way it is easy to represent different fusion techniques. The components responsible for the fusion stages are the following:

- DAI-DAO (Data In – Data Out)
- DAI-FEO (Data In – Feature Out)
- FEI-FEO (Feature In – Feature Out)
- FEI-DEO (Feature In – Decision Out)
- DEI-DEO (Decision In – Decision Out)

where DAI-DAO corresponds to Low Level Fusion, FEI-FEO to Medium Level Fusion, DEI-DEO to High Level Fusion, and DAI-FEO and FEI-DEO are included in Multilevel Fusion. The main contribution of Dasarathy's classification is that it specifies the abstraction level of both the input and the output of a fusion process, avoiding possible ambiguities. However, this functional model refers to a data driven process [44], where an overall systemic view is not provided and the user role cannot be accommodated. In [74] a mapping between Dasarathy's functional model and JDL model has been provided.

OODA loop model

Another model for IF, mainly developed in the military field, is the OODA (Observe-Orient-Decide-Act) loop. The aim of the model is to enable faster decisions by identifying both your own decisions and your opponent's ones, in order to act before your opponent. Despite the fact that the OODA-loop is quite simplistic, it is the most accepted decision making process model used within information fusion. The four activities considered in the process are the following (see Figure 7):

- *Observe*: the environment, in order to detect an opponent;
- *Orient*: position yourself in the environment, in a good place for the next step;
- *Decide*: make a decision, based on previous stages;
- *Act*: perform the decision.

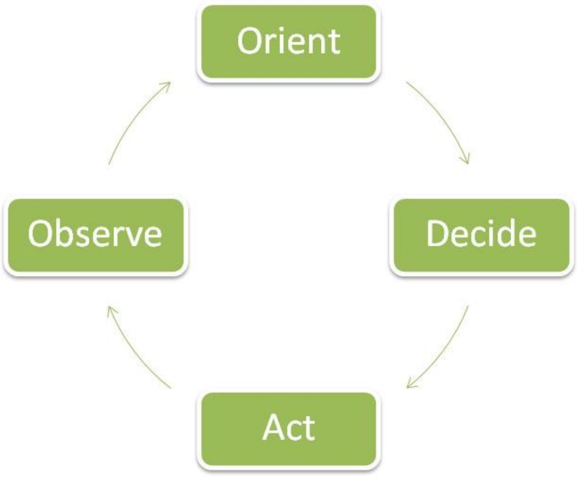


Figure 7 – OODA loop

The model illustrates the ultimate goal of a decision maker, taking the right decision within the minimum time, where speed is a condition for winning. Although the OODA loop has its origins in the military domain, it focuses more in general in the human decision process. Besides, the only military-specific term is "*orient*", so by replacing this term with "*interpret*" (to represent the concept of situation understanding), the model becomes more generic.

Formally, an extension of the original OODA loop model in order to improve the capacity to represent dynamic and complex situations by a modular approach, is proposed with the M-OODA (modular OODA) loop [70]. It consists of four goal-oriented modules (more generic than the original four activities):

- *Data gathering* (Observe);
- *Situation understanding* (Orient);
- *Action selection* (Decide);
- *Action implementation* (Act).

In addition, each module is structured around three components: Process, State, and Control. The M-OODA loop incorporates explicit control elements within and across modules enabling a bidirectional data/information flow between modules. It also includes a feedback loop within each module. Finally, it provides a basic architecture for modeling a variety of team (rather than individual) decision-making, differently from the OODA loop.

Object Oriented reference model

The object oriented reference model represents a formal approach to fusion system design and it shows the role of the psychology of the human-computer interface in the system design process [49]. In fact, with this model the human capabilities can naturally find their space. In particular, the model does not specify fusion tasks or activities, however it provides a set of roles and specifies the relationships among them. The identified roles are:

- *Actor*: responsible for the interaction with the world, collecting information and acting on the environment;
- *Perceiver*: once information is gathered, the perceiver assesses such information providing a contextualized analysis to the director;
- *Director*: based on the analysis provided by the perceiver, the director builds an action plan specifying the system's goals;

- *Manager*: the manager controls the actors to execute the plans formulated by the director.

The proposed model perspective is such that human and computer objects are not distinct.

Waterfall Model

The Waterfall model is an example of hierarchical architecture, described in [45]. The flow of data operates from the data level to the decision-making level, where the sensor system module (Level 1) is continuously updated with feedback from the decision-making module (Level 3). The intermediate level is responsible for the pattern processing. The three levels are described as follows:

- Level 1 transforms raw data to provide the required information about the environment;
- Level 2 is composed of feature extraction and fusion in order to obtain an inference about the data. The output of this level is a list of estimates with associated probabilities;
- Level 3 relates objects to events, according to the information that has been gathered, the available libraries and databases, and the human interaction.

A detailed schema of the model is proposed in [28].

Omnibus Process Model

This model draws together several models, taking their advantages and overcoming some of their disadvantages, presenting a general taxonomy to capture the IF process [8]. The models involved are: JDL, OODA loop, DFD, and the waterfall model. Omnibus process model includes a dual perspective: system and task oriented. The decision making is considered as a computational process. It can be considered as Level 4 of the JDL model, and as the Decide phase of the OODA loop.

1.3 Event correlation techniques

Whatever the monitoring solution (through a single sensor, a multimodal sensing, an integrated surveillance system, or a PSIM system) and the IF strategy, in the described context the objective is to capture events that occur in the monitored environment. At the same time, the event correlation is aimed at capturing all the further events of interest within the application domain, in such a way to undertake ad-hoc actions or procedures (sequences of actions) to manage the situations corresponding to these events. That is the main reason why, in such applications, is common to refer to an event-driven approach to design real-time detection systems. Many practical applications require to react to complex event patterns, rather than single event occurrences. To simplify design and analysis of such reactive systems, it is useful to separate the detection mechanism of complex event patterns from the rest of the application logic, e.g. in charge of a PSIM system.

Complex Event Processing (CEP) is one of the main field of research dealing with the handling of events. It is also closely linked to the Event-Driven Architecture (EDA) paradigm, which concerns detection, consumption of, and reaction to events. CEP is on the top and it filters, matches, and aggregates events into higher-level events [69]. CEP is aimed at the processing of information streams (associated to the occurred events) to infer complex events patterns corresponding to more complicated situations. In the context we move, each event is associated to a meaningful change of state in the monitored environment. The information to combine in real-time comes from multiple sensing sources and the complex events represent threat scenarios, to which one should respond, quickly and accurately.

The topic of correlation is widespread in many fields, like data networks (e.g. to detect faults or security-related attacks) and active databases (e.g. to enable the recognition of complex combinations of events). According to the application, specific complex event patterns should be matched against the streams of occurred events during the run-time of the system.

Regarding the physical security of infrastructures, we can define the event correlation task as an interpretation procedure to confer a new meaning (indicative of possible threats) to a set of simple events (detected by each sensing system) happened within a predefined time interval. The procedure could range from trivial event compression, filtering, counting to complex pattern matching. Correlation can also consider temporal relations between events, while the event clustering allows the creation of complex patterns e.g. by means of boolean operators. The single terms in the pattern could be primary or higher-level events (i.e. generated by the correlation process) [15]. The main event correlation approaches for situation recognition are the following:

- Rule-based

  It is based on "if-then" rules, which represent the specific knowledge, relevant for the application. A well-formed rule consists of a prerequisite (or a set of prerequisites), which must be satisfied to apply the rule, linked to an action to be performed if the rule is applied. A rule-based system has an inference engine (to define the order with which the applicable rules will be executed), a knowledge base (including the set of all the rules), and a working memory (containing the data about the current monitored situation). The rules matching can be exact or partial, if all the prerequisites should be satisfied or not, before executing the action. In several contexts, such rules are often named Event Condition Action (ECA) rules. From a certain point of view, majority voting methods also belong to rule-based category, since they achieve a final "opinion" on a specific topic, according to a majority rule (absolute or relative) applied on a set of opinions, i.e. votes [6] [65]. Typically, rule-based systems have a stateful and offline execution mode. In the existing literature, the approach has been successfully used, for example, in face detection, human tracking and person identification. However, from the viewpoint of this thesis, we assume that each sensing system (e.g. VMS, ACS, and so on) takes already in charge all these tasks at a lower abstraction level.

- State transition based

  The typical use of this approach is aimed at recognition of patterns in sequences of symbolic input. Finite State Machine (FSM) and Petri Net belong to this category. FSM is characterized by a finite set of states, a pre-defined starting state, an alphabet of possible input and output events, and a state transition function which determines the next state and the possible output (optional) for each state and input event [71]. An application example, where complex event detection is implemented using automata, is the Ode approach for active databases [39]. In this case, management and manipulation of the database is performed by means of an object oriented language.

  A Petri net is useful to represent causal dependencies and concurrent processes [71]. It is a directed bipartite graph, in which the nodes represent transitions (symbolized by bars), i.e. events that may occur, and places (symbolized by circles), i.e. conditions. The directed arcs (symbolized by arrows) run from a place to a transition or vice versa (never between places or between transitions). A Petri net may contain a certain number of tokens. Only with a sufficient number of tokens at the start of all input arcs, each transition is enabled. Once executed, it places these tokens at the end of all output arcs. On this formalism is based the composite event detection of Samos approach for an Object Oriented Data-Base Management System (OODBMS) [38]. In this case, each primitive event is represented by a Petri net place. The event occurrences are entered as individual tokens. Further net places and transitions represent complex event expressions. A different approach is based on the semantics of the Snoop event algebra, including a set of operators [17]. Its concepts have been implemented in a prototype called Sentinel [18] for an active object-oriented database. Composite event detection is graph-based: simple event occurrences enter the bottom nodes and flow upwards through the graph, being joined into composite event occurrences. Besides, ad-hoc consumption modes of simple event occurrences augment the semantics of composite events. In [5] the implementation of an event detection engine for the Web, that detects composite events specified by expressions of an illustrative sublanguage of the Snoop event algebra, is presented.

- Classification-based

  The approach is used to obtain a decision, classifying an observation into one of the pre-defined classes. Support Vector Machine (SVM) is a popular method for classification tasks, also in the domain of multimedia surveillance. It is a binary classifier exploiting a supervised learning method. In order to detect high level semantic concepts, the input of a SVM could be a vector including all the low level features, e.g. related to video, audio, and so on [2]. A further classification-based approach is the Bayesian network (BN, also named belief network). It is a directed acyclic graph which models probabilistic relations, e.g. between the threats and the primitive events detected by sensors. Given the event occurrences, a BN can compute the probabilities of the threats. The approach is widely used in multimedia analysis [6][7]. However, it requires well defined a-priory and conditional probabilities of the hypothesis (e.g. about the threats and related triggering events) to be effective. Finally, Artificial Neural Network (ANN) is another approach adopted for event correlation. The idea is to reproduce with an artificial model the functioning of the human brain. ANN is made up by nodes, which perform operations on weighted inputs to get an output, possibly used as input for further nodes [51]. The processing can be of several typologies (mathematical, temporal, etc.). Ad hoc procedures can be used for adapting the weight dynamically. ANN is a non-linear black box that can be trained to solve complex and high dimensional problems, however the selection of a proper network architecture (for the specific application) can be difficult and the training can be time-consuming [6].

The logic and temporal aspects of event correlation are the main concerns addressed in Chapter 3, where a graph-based event detection is described.

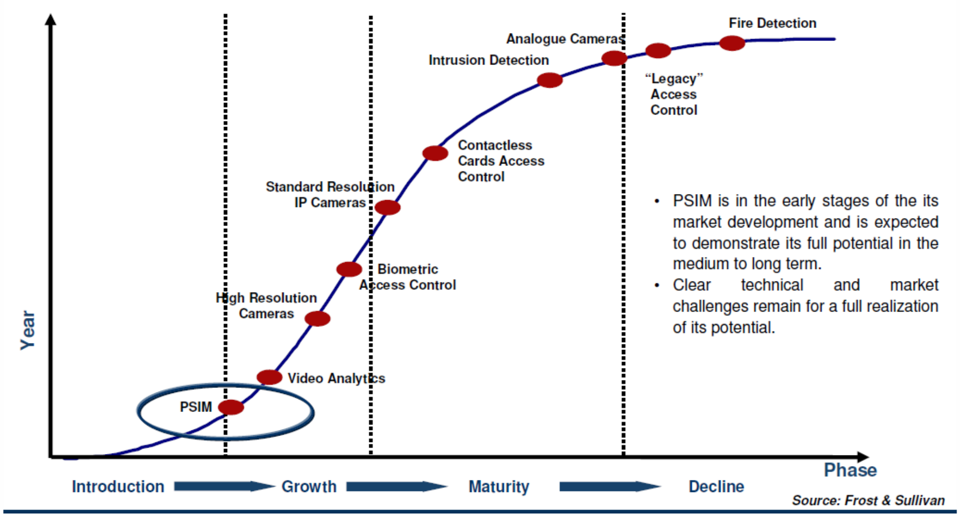
1.4 Open issues

In all the activities concerning the described context, Information Technology (IT) plays an important role, since it enables new and effective means to mitigate risks, providing early warning of threats and improving the response to disasters of various severity. As such, IT has an impact also in increasing CIs resilience. In fact, sensor-based technologies for detecting meaningful events can help in preventing unwilled situations. Traditionally, at least (digital) video-surveillance and intrusion detection technologies have been employed. However, for an enhanced early warning and situation awareness, the traditional technologies are not enough.

Often, the management of the security-related events is fragmented. Each event is treated separately. And many times there is a lack of an effective information sharing. The key to overcome those limitations is to synthesize data from multiple alerting systems and physical sensors. The use of distributed and heterogeneous sensorial subsystems (encouraged by the development of novel low-cost sensing devices) and their integration, can lead to several levels of event correlation. A low level integration allows for the development of multi-modal approaches for monitoring and surveillance activities. Such a solution aims at providing advanced event detection capabilities and/or at improving detection reliability. Some of these benefits are counteracted by additional problems, like harder calibration and correct management of available cameras, for example in multi-camera systems. In many installations, multiple sensing systems are integrated at junction-box level. The result is a more immediate and reliable data correlation, however the integration may be more difficult for the different representations (and semantic levels) of the combined modalities (e.g. audio and video), and the scalability of the solution can be limited with respect to the needs. In fact, many infrastructures (especially transportation systems) can be spread through hundreds of kilometres. In addition to that, they can require thousands of cameras and other sensors, which makes human-based surveillance unfeasible. Manual analysis of video as well as diverse sensor alarms (which can be false) is labour intensive, fatiguing, and prone to errors. Additionally, psychophysical research indicates that there are severe limitations in the ability of humans to monitor simultaneous signals. Thus, it is clear that there is a fundamental contradiction between the current surveillance model and human surveillance capabilities. Therefore, software-aided real-time video analytics considerably alleviates the human constraints, which currently are the main handicap for analyzing continuous surveillance data [16]. However, though video-analytics may not be considered as a novel development (in fact, Computer Vision research has been active since early 80s), recent experiences using state-of-the-art systems reported low performance in terms of false alarms and missed detections. Therefore, the redundancy and diversity of sensing technology is essential to build effective surveillance systems. That increases the number of sensing devices and - consequently - of the alarms to be integrated and managed.

High level information aggregation is the key to develop the next generation of security management systems, the so called PSIM (Physical Security Information Management) systems. They help in integrating security devices, in improving detection efficiency and effectiveness, and in enhancing the overall situation awareness and management. One of the key factors to achieve those results is the presentation of all the relevant information into a single view, in order to provide essential decision support features.

In this thesis we have considered IF strategies and event correlation capabilities as the main resources to address these issues. Regarding IF, the main open issue is to create a convergence with the Decision Support (DS) research field. IF and DS areas has been developed independently from each other. However, there is the real need to provide the user's perspective in the IFS design, and to exploit IF in the DSS design as a technique to support and improve decision-making [60]. At the same time, event correlation capabilities should help in providing data analysis and interpretation in real-time, which represents an added value of the PSIM systems with respect to the traditional security management systems (like VMS, ACS, etc.). Of course, the lack of advanced applications in PSIM is motivated by the still missing light, efficient, and easy-to-use correlation approaches.



- PSIM is in the early stages of the its market development and is expected to demonstrate its full potential in the medium to long term.
- Clear technical and market challenges remain for a full realization of its potential.

Source: Frost & Sullivan

Figure 8 – PSIM life cycle analysis

Nowadays the effective exploitation of all these potentialities represents a way to overcome the main obstacles to the PSIM deployment in the years to come. They are mainly due to the lack of understanding of what PSIM is/does, the high costs, and the partial overlap with other tradition systems (like the VMS) [46]. It is worth noting that doubts and limitations regarding PSIM are also the consequence of the early stages of its development (see Figure 8), which will demonstrate its full potential in the medium to long term [36].

In the context described above, the novel research must aim at detecting threats scenarios as early as possible, providing superior situation awareness and decision support to quickly – possibly automatically – activate response-and-recovery strategies. That can be achieved by means of ad-hoc information fusion and event correlation techniques. All those aspects contribute to what we define “augmented surveillance”: an integrated concept, including technologies, capabilities and functions, which we believe will be one of the most challenging paradigm of the CIP (and not only) research in the future. Chapter 2 presents a more detailed description of such a paradigm, while Chapter 3 describes an advanced framework for event correlation in PSIM. Chapter 4 and Chapter 5 introduce the further developments of the framework aimed at improving detection effectiveness and efficiency. Real applications in a more specific domain and the overall integration with an existing PSIM system are finally included in Chapter 6 and Chapter 7.

Chapter 2

Towards an augmented surveillance paradigm

2.1 Basic concepts and requirements

This section describes a paradigm aimed at the physical protection of assets in threat scenarios. The basic idea is to collect and exploit all the means, techniques, methods, and approaches already available (and briefly described in the previous sections), in order to obtain an integrated layered platform easy to adapt to the specific application domain.

The key to achieve “augmented surveillance” is to combine and synthesize data from multiple and distributed sensorial subsystems, at different levels. First of all, technological issues, techniques and methodologies of protection, and overall fusion strategies should be viewed in a cohesive way. Furthermore, ad-hoc information integration and management should allow PSIM systems for reporting the available (or inferred) information into a single integrated view, to increase effectiveness and efficiency in decision support. The new PSIM generation can significantly contribute to improvements not only in threat detection, but in several directions, such as:

- deterrence: to discourage the adversary from acting;
- minimization: to mitigate the effects of attacks;
- response: to enable operators to counteract the attack and to protect assets and persons;
- recovery: to enable the system to resume normal operations.

The proposed paradigm is aimed at reaching advanced early warning capabilities, inside a general context of enhanced Situation Awareness (SA). It aims at aiding decision makers in obtaining a greater knowledge of events, factors, and variables affecting a certain environment. According to one of the first and most widely accepted definitions, SA is the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future [26].

SA includes also the important concept of “situation recognition”, which aims at identifying a-priori defined situation patterns within an information flow, in order to support decision makers allowing them to focus only on the most relevant aspects. Situation recognition can be considered as a pattern matching problem, where patterns represent situations of interest. In such a task, often the issue is not the lack of information, but finding the information needed when needed [27]. That requires the use of computer-based support systems, since the human operators may not be able to analyze all information properly and timely. Further background concepts specifically related to multimedia surveillance and monitoring systems are provided in [7].

From the physical security viewpoint, the a-priori defined situations of interest include the threat scenarios, which are identified during the phase of Risk Analysis (RA), performed for the infrastructure or the assets to be protected. They are typically composed by sequences of actions used by attackers to reach their objective. The preliminary stage of RA regards the adoption of rigorous and systematic approaches to model possibly complex threats. The aim is to identify and to model, using a certain formalism, the possible modes of attack (i.e. the threat scenarios). Each threat is typically associated to a risk index to obtain a quantitative or qualitative classification. This is essential to define a priority in the adoption of countermeasures and protection mechanisms. Therefore, for all the subsequent stages (selection of detectors, system deployment, definition of the IF strategy, etc.) an a-priori knowledge about the possible threat scenarios can be assumed.

The increasing need for correlating large heterogeneous information to provide greater SA and early warning should fulfil several requirements. First of all, given a known threat and assuming that:

- all the means (i.e. sensors and devices) for threat detection are available;

- each device works properly, i.e. it is in the condition to correctly detect a threats trace (data, status, event, etc.);

the detection of the threat and the report of related alarm to the operators should be assured. In other words, given a problem, the existence of a “solution” should be assured. This affects the overall approach (e.g. correlation techniques and related models) which lies behind the detection mechanism implemented by the surveillance systems.

In addition to that, a quick data processing is required. Therefore the logic which rules the behavior of the surveillance system should be based on models sharing the requirement of (soft) real-time solvability.

The requirements above mentioned should be fulfilled also with a certain level of reliability. In fact, the surveillance systems should assure as much as possible a high POD (Probability Of Detection) and a low FAR (False Alarm Rate): those parameters determine the effectiveness and the viability of the system, respectively (see section 2.2).

These considerations give useful indications in defining the right approach to face the detection problem. For example, an Artificial Neural Network (ANN) represents a possible approach to solve large and computationally demanding problems. It is usually used to model complex relationships between inputs and outputs or to find patterns in data. However, ANN is also a non-linear black-box, which require a training phase. Therefore, it is necessary to reason on more specific constraints:

a. to assure the predictability and repeatability of the behaviour of the system (e.g. within an ANN, it is not possible to describe the stored Knowledge Base, KB);

b. to overcome the critical concern of the learning phase, that can require sophisticated training techniques, long computation time, and a large set of examples;

c. to fulfil the requirement related to the existence of a “solution” (e.g. ANN cannot assure a solution to the posed problem, since the existence of a learning algorithm which converges is not guaranteed).

The above statements suggest using expert systems, which are not "intelligent" in the usual sense of the word, i.e. in a creative way. While the deductions of expert systems are constrained by the stored KB, they can process a large amount of data very quickly, taking into account many “rules” and details that human experts cannot do. One limit of this approach is that the completeness of the KB depends on the effectiveness and quality of the RA activity: threat scenarios that are not identified and translated into a model will never be detected.

This issue leads to a further reasoning on the capabilities to be provided by modern PSIM systems featuring a certain level of ‘augmented surveillance’. The model-based threat detection approach, in fact, should also assure a certain tolerance to imperfect modeling (due to human faults) on the one hand, and to missed detections (due to device faults) on the other. The set of possible solutions includes techniques of pattern matching and similarity analysis, in order to recognize new and not (perfectly) modeled threats. The techniques based on similarity between patterns are not new, in particular in the field of computer network intrusion detection systems. They are often based on the following solution: if an alert, which is the known consequence of a forerunning event, is received and the forerunning event has not been detected, then the missed event can be identified [21]. The limit of this solution lies in the fact that it cannot cope with missing events that are not linked to other events. A more effective technique should not assume that direct cause/consequence relationships exist between detected and missing events. For example, this is possible with solutions based on ad-hoc metrics. With respect to traditional approaches of infrastructure surveillance, such a technique allows for earlier as well as more robust and straightforward detection of complex threat scenarios. It does not require further modeling efforts, since threat scenarios do not need to evolve completely to provide a warning: operators may receive a warning level which is somehow proportional to the similarity index. These quantitative indications about unfolding threats can effectively help operators to quickly undertake appropriate countermeasures [32].

Innovative approaches for the design of distributed surveillance systems should aim at adding interactivity and adaptability capabilities, fulfilling the constraint of a preventive and manageable mode of reaction. Recent studies on Cognitive Systems (CS) help in reaching this goal. As matter of fact, automatic surveillance systems are required to emulate the cognitive capabilities of human operators in detecting and assessing possible threats. In particular, this kind of approach is increasingly used in the field of intelligent video surveillance [54], not only to understand complex activities occurring within a video stream, but also to learn from them in such a way to build a knowledge base automatically adapted to the specific environment. Recently, the cognitive paradigm has been also applied to the field of CIP [19], in order to provide a more comprehensive situation awareness. Methods for the representation and the organization of knowledge and for the learning from experience allow a system to evaluate the state evolution and to predict near future events. The emerging concept of cognitive surveillance, based on this kind of researches, aims at providing these capabilities also by means of the cooperation with human agents to perform corrective actions. Regarding the cognitive cycle, discussions about the correctness and suitability of the semantic descriptions of events of interest (depending on the domain) are provided in [29].

2.2 Viability and effectiveness

The human management of critical situations involving many simultaneous events is a very delicate task, which can be error prone. Integrated surveillance systems are necessary to allow the human supervision of a large number of sensors, devices, or cameras positioned inside the environment to be monitored. These systems allow to call the attention of the operators only and anytime an alarm is detected, trying to make surveillance independent from their attention level. Generally speaking, the concern is related to a quick and effective management of possible large amounts of data (e.g. events, alarms, etc.). Therefore the first challenging goal is to support and strengthen the human capabilities without replacing them. The main motivation relies in the need for taking into account many details that a human could ignore, miss or forget.

Since a few human operators are usually employed in security surveillance, human-factors need to be carefully addressed, including cognitive ergonomics in human-machine interaction. In fact, many critical tasks are under the responsibility of human operators, that cannot manage a great amount of surveillance streams in real time. Hence, it is necessary to find the best trade-off between the tendency of sensing subsystems to produce a large amount of data (events, warnings, alerts) and the limited human capabilities. In the field of Human Factors and Ergonomics (HFE), it is widely believed that highly non-stimulating and repetitive activities make human surveillance very difficult [73]. This is especially true in multimedia surveillance systems, where the operators may monitor a wide area, through a large number of sensors producing many events, warnings and alerts.

Furthermore, modern surveillance systems typically support the undertaking of countermeasures, whose activation can be fully automatic (independent from human intervention) or partially automatic (based on human discrimination, e.g. by manual confirmation of detected alarms). The choice of the response mode may depend on the kind of countermeasure, but also on timeliness requirements and on the alarm trustworthiness.

Obviously, alarm systems should only detect situations that actually represent a threat. However, intelligent sensing systems may generate unnecessary warnings, which can be classified as false alarms or nuisance alarms. Therefore, with regard to the triggering of countermeasures, it is very important to take into account and to control the rate of these alarms.

False alarms are due to events that should not cause an alarm, while nuisance alarms are generated when a legitimate cause occurs, but the related alarm activation is inconvenient. As an example, nuisance alarms occur when maintenance staff enters restricted areas without prior identification. False and nuisance alarms can have a significant impact on:

- operational efficiency, due to the time wasted in evaluating and dismissing unnecessary alarms as well as in the possible de-activation of automatic countermeasures;
- vigilance level and response time, since when a large number of alarms are false, operators tend not to trust them.

In other words, if the probability of detection determines system effectiveness, false and nuisance alarm rates significantly influence its operational viability and efficiency. Therefore, it is essential to identify reasonable goals for detection and false alarm rates, and then to determine the methods to achieve them.

Considering the joint automation-human performance and in particular how the level of automation unreliability affects human performance – a research revealed that alarm systems should have a reliability factor of 70% [82]. At a reliability level below this threshold, the automation can be considered worse than no automation at all, nullifying its practical usefulness. The analysis also showed that performance was more strongly affected by reliability in high workload conditions, which are critical in the context of surveillance and supervision.

According to this result, when more than 30% of alarms are false or nuisance, operators: waste time in discarding alarms; ignore or respond slowly to real events; lose confidence in the surveillance system. This aspect is tied to the adverse effects of false alarms on human behavior. The rule mentioned above is important to establish a minimum level of reliability in order to achieve a viable system.

Many approaches address the issue of nuisance alarms in a retroactive way, in particular employing self-learning engines to filter out unnecessary alarms based on human feedback. With this capability, the system can adapt to the specific conditions of each installation and learn to recognize events which are cause of false and nuisance alarms. The difficulties of this approach lie in the self-learning process, recognition strategy and duration.

One obvious method to improve detection rates is to increase the sensitivity of the sensing subsystems; however, this will also increase the number of false alarms. Following the usability criterion mentioned above, the surveillance system (and/or its HMIs) should be optimized in such a way to get the highest detection rate with no more than 30 discardable alarms out of any 100 generated (on average). An improvement of alarm trustworthiness as well as system resilience can be achieved by exploiting redundancy and diversity in sensors displacement and technologies.

2.3 Practical applications

In this section, a practical application of augmented surveillance is proposed. One of the basic needs is to frame in a cohesive way technologies, techniques, capabilities, and features available for distributed intelligent monitoring in order to merge the two areas of Information Fusion (IF) and Decision Support (DS). More in detail, in the field of IF there is a lack of the research focusing on the user's decision making process embedded in an information fusion system, that is essential to fully take advantage of its benefits [60]. In order to represent the overall process describing how information transforms from sensor data to information which a user can use for decision making, a general model is proposed. It is based on the JDL model (see section 1.2.2), because it highlights three important aspects that are reflected in the domains of interest:

a. The possible need for pre-processing raw data coming from sensors. This aspect depends on the level of heterogeneity of the sensing subsystems (ranging from temperature sensors to intelligent cameras) and on the semantic level of the information provided by them. The issue could be addressed by the subsystems, by the PSIM which integrates them (by means of an “Adaption and Connectivity” layer, as shown in Figure 3), or even before information fusion processing.

b. The identification of different levels of capabilities. The data combination from sources is aimed at evaluating: i) states, attributes or identities of the monitored entities; ii) mutual relationships between monitored entities and surrounding environment; iii) future states and projections starting from recognized situations, in order to assess threats, risks and possible impacts.

c. The need for performing a constant refinement process, which can be automatic, user-driven or hybrid. In fact, the user can effectively contribute to this process to complement the information fusion system [10].

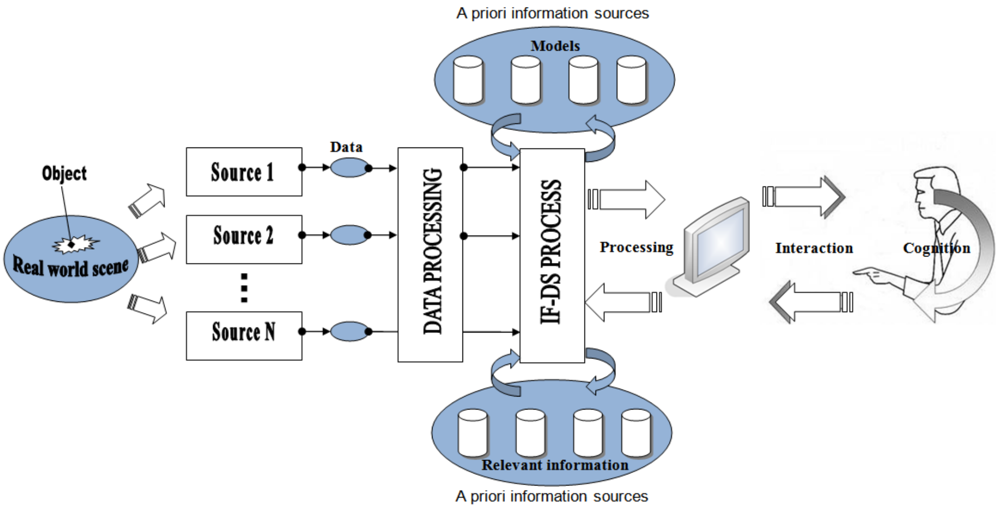


Figure 9 – Overall data/information management process

An overall process for handling data from multiple sources is represented in Figure 9. In the figure, each source may represent a single sensor (regardless of its ‘intelligence’), a complex – possibly multimodal – sensing subsystem, or a human agent. Depending on the type of outputs, a preliminary processing may be required or not. The figure remarks the important aspect of considering the IF and DS processes as a whole. Furthermore, it takes into account both the knowledge base (represented by detection models) and the relevant information data-bases (e.g. the ones used to store the user feedbacks) as information sources, which can be updated during the fusion process. Finally, the interaction with the final user allows for understanding situations, recognize emerging trends, undertake decisions and countermeasures.

From the application point of view, the augmented surveillance paradigm could be implemented by an integrated framework addressing surveillance-related information at different levels; its PSIM capabilities should include:

- the integration and interfacing with several heterogeneous sensing platforms, in such a way to solve the problem of the preliminary data processing;
- the detection of relevant events occurring in all the monitored locations;
- the warning and alarm reporting to operators in control centers, in order to support the emergency procedures and/or to activate automatic reactions.

Since the PSIM may generate a large amount of alarms which could overwhelm the personnel in charge of reacting to suspicious events, event correlation capabilities need to be integrated into the framework in order to lower the false alarm rate and to improve threat detection reliability. The problem of event correlation has been largely studied in the scientific literature, and a wide class of potential solutions have been defined. Nevertheless, those results have been widely studied and applied to domains not related to physical security, e.g. to develop intrusion detection in computer networks. In physical security applications, the capabilities of legacy systems are very limited in data analysis and interpretation, and hence in real-time prevention and reaction. The lack of advanced approaches in PSIM may be motivated by the still missing light and efficient approaches to the recognition of evolving situations based on a-priori knowledge of threat patterns, to be easily updated by the human operators:

such an objective is far from being trivial to achieve. However, it is increasingly important to achieve early warning and situation awareness in domains where a large number of dynamic objects are engaged in complex spatial-temporal relations.

In the assumptions that threat scenarios can be decomposed in a set of basic steps executed in a predictable sequence, model-based logical, spatial and temporal correlation of detector outputs can be used to recognize event patterns indicating possible threats. Ideally, in order to recognize (partially) unknown threats, the detection engine should be resilient to human faults in scenario modeling and sensor faults in detecting events. One possibility which has been recently researched is to consider heuristics like similarity analysis with known event patterns (see e.g. references [13] and [32]).

Chapter 3

The DETECT framework

3.1 General description

The best way to face threats is to stop them before they can cause serious consequences. Unfortunately, visual surveillance of video streams and sensor alarms provided by traditional physical security systems does not allow human operators for a satisfactory situational awareness when the sequence of events is large, heterogeneous, geographically distributed and rapidly evolving. Therefore, human operators may not be able to recognize sequences of events which are indicative of possible threats due to their limited alert threshold and knowledge base. Furthermore, operators can be unable to guide and coordinate alarm responses or emergency interventions (in particular in case of simultaneous critical events) if they are not precisely aware of what is happening or has happened. In order to cope with these issues, early warning and decision support systems should be adopted to face physical security threats, by means of heterogeneous distributed sensing subsystems.

The heterogeneity of technologies and data requires integration and analysis at different levels. Assuming to solve the integration issue with a PSIM system (which is, for example, interfaced with each sensing subsystem through proper SDKs and APIs), there is the need for an on-line reasoning about the events captured by sensors, in order to early detect and properly manage security threats. The possible availability of heterogeneous and redundant information allows for the correlation of basic events in order to increase the overall probability of detection (POD), decrease the false alarm rate (FAR), warn the human operators about suspicious situations, and even enable the automatic trigger of adequate countermeasures.

This section describes the motivation, the working principles and the architecture of DETECT (DEcision Triggering Event Composer & Tracker) [35], a framework aimed at the automatic detection of physical security threats, possibly before they evolve to disastrous consequences. In fact, non trivial threat scenarios are made up by a set of basic steps which have to be executed in a predictable sequence and with possible variants. Such scenarios must be precisely identified during the important phase of Risk Analysis, to be performed on the infrastructures that one intends to protect [52]. DETECT operates by performing a model-based logical, spatial and temporal correlation of basic events detected by each monitoring subsystem, in order to "sniff" sequence of events which indicate – as early as possible – the likelihood of threats. In order to achieve this aim, DETECT is based on a detection engine which is able to reason about heterogeneous data, implementing the concept of "fusion" trough event correlation. The framework can be interfaced with existing PSIM systems, in order to effectively enhance the situation awareness and improve the decision making process of human operators. It can serve as an early warning tool, since it can alert the operators about the likelihood and nature of the threat, and consequently even to support the (automatic or manual) activation of adequate countermeasures for emergency/crisis management. As such, it may allow for a quicker and more focused response to threat scenarios, possibly before they can evolve. This feature represents a crucial point in the context of the human management of critical situations (in particular if it involves many simultaneous events). In fact it is a very delicate task, but also error prone as well as subject to forced inhibition. In addition to that, the correlation among basic events detected by diverse redundant sensors allows to lower the false alarm rate of the security system, thus improving the overall reliability of the security system.

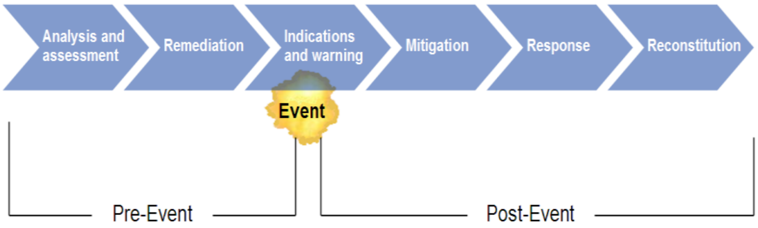


Figure 10 – CIP life-cycle

Finally, with particular reference to CIP application domain, DETECT is mainly located at the third stage (i.e. "Indications and warning") of the CIP life-cycle reported in Figure 10. However, thanks to the overall integration with a PSIM system and to the developments described in Chapter 4 and Chapter 5, the framework can involve also the other pre-event and post-event stages.

3.2 Event description language

Threats scenarios are described in DETECT using a specific Event Description Language (EDL) and stored in a Scenario Repository. In this way it is possible to store permanently all scenario features in an interoperable format (i.e. XML[3]). At the same time, the Event History database contains the list of basic events detected by the sensing devices. A high level architecture of the framework is depicted in Figure 11.

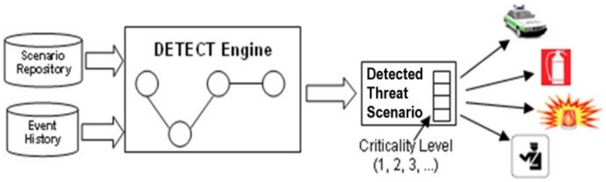


Figure 11 – The DETECT framework

The Detection Engine needs to recognize combination of events, bound each other with appropriate operators in order to form composite events of any complexity. The EDL of DETECT is derived from the Snoop event algebra [17], so let us define some basic concepts accordingly.

[3] XML (eXtended Markup Language) Metadata Interchange.

A *threat scenario* consists of a set of basic events (detected by the sensing devices), which occur in a predictable sequence with possible variants.

An *event* is a happening that occurs at some location and at some point in time. In this context, events are related to sensor data (e.g. motion detected by a camera, intrusion detected by a volumetric sensor, etc.).

Events are classified as *primitive events* and *composite events.* A primitive event is a condition on a specific sensor which is associated some parameters (i.e. event identifier, time of occurrence, etc). Primitive events represent the basic events mentioned previously. All the occurrences of primitive events are stored in the Event History database, whose schema includes at least the information indicated in Table 1.

Field Name	Field Description	Field format (example)
IDev	Event Identifier	E*x* (e.g. E8)
IDs	Sensor Identifier	S*x* (e.g. S4)
Tp	Timestamp	*yyyy-mm-dd hh:mm:ss* (e.g. 2010-10-01 23:56:09)

Table 1 – Parameters associated to a primitive event occurrence

Since there is the need for specifying complex patterns of events, it is important to define an appropriate set of *Operators*. They allow to express relationships between primitive events and thus to combine them in a meaningful way. Therefore, a composite event is a combination of primitive events by means of proper operators (logical, temporal, etc.). Formally an event E (either primitive or composite) is a function from the time domain onto the boolean values, *True* and *False* [18]:

E: $T \rightarrow$ {True, False}, given by:

$$E\,(t) = \begin{cases} \textit{True}\text{, if E occurs at time t} \\ \textit{False}\text{, otherwise} \end{cases}$$

The basic assumption of considering a boolean function is quite general, since different events can be associated to a continuous sensor output, according to a set of specified thresholds. Furthermore, negate conditions (!E) can be used when there is the need for checking that an event is no longer occurring. This allows considering both instantaneous ("occurs" = "has occurred") and continuous ("occurs" = "is occurring") events. However, in order to simplify EDL syntax, negate conditions on events can be substituted by *complementary events*. An event E_c is complementary to E when: $E_c \Rightarrow !E$.

Each event is denoted by an *event expression*, whose complexity grows with the number of involved events. Given the expressions E_1, E_2, …, E_n, every application on them through any operator is still an expression. Each event expression is represented by an *event tree*, where primitive events are at the leaves, while internal nodes represent the operators [33].

The EDL considers the following operators: OR, AND, ANY, SEQ. As an example, Figure 12 shows an Event Tree for representing an event expression. Leaf nodes E1, E2 and E3 represent primitive events and internal nodes represent the "AND" and "OR" composite events. The whole event tree represents the composite event "(E1 AND E2) OR E3".

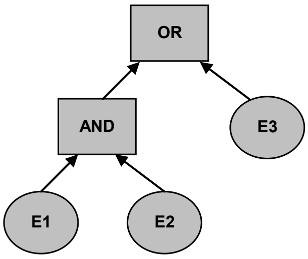


Figure 12 – An example of event tree

In summary, threat scenarios are identified during the phase of Risk Analysis, performed for the infrastructure to be protected, while the primitive events are the ones detectable by the sensing devices installed on the field. The latter are related to sensor data variables (i.e. variable *x* greater than a fixed threshold, variable *y* in a fixed range, etc.). Using operators, it is possible to compose more complex events, which represent the threat scenarios indentified during the risk analysis.

The semantics of the Snoop operators is as follows:

- *OR*: disjunction of two events E_1 and E_2, denoted with (E_1 *OR* E_2). It occurs when at least one of its components occurs.
- *AND*: conjunction of two events E_1 and E_2, denoted with (E_1 *AND* E_2). It occurs when both events occur (the temporal sequence is ignored).
- *ANY*: a composite event, denoted with *ANY*(m, E_1, E_2,…, E_n), where $m \leq n$. It occurs when m out of n distinct events specified in the expression occur (the temporal sequence is ignored).
- *SEQ*: sequence of two events E_1 and E_2, denoted with (E_1 *SEQ* E_2). It occurs when E_2 occurs provided that E_1 has already occurred. This means that the time of occurrence of E_1 has to be less than the time of occurrence of E_2.

The sequence operator is used to define composite events when the order of its component events is relevant. Another way to take into account the time in the event correlation is by exploiting explicit *temporal constraints*. They can be specified on operators, to restrict the time validity of logic correlations. In fact, the latter could lose meaningfulness when the time interval between component events exceeds a certain threshold. Therefore, the definition of temporal constraints has the aim of setting a validity interval for the composite event. Such constraints (e.g. expressed in seconds) can be added to any logical operator in the formal expression used for event description. For instance, let us assume that in the composite event E = (E_1 AND E_2) the time interval between the occurrence of primitive events E_1 and E_2 must be at most T. The formal expression of the event E is modified by adding the temporal constraint [T] as follows:

$$(E_1 \text{ AND } E_2) [T] = \mathit{True}$$

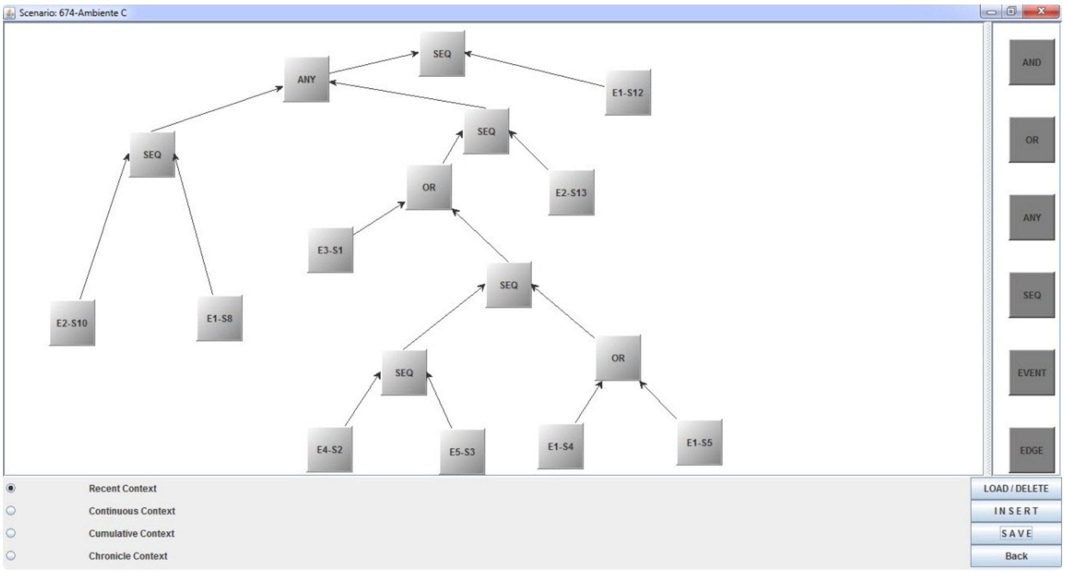
$$\Leftrightarrow$$

$$\exists\, t_1 < t \mid (E_1(t) \wedge E_2(t_1) \vee E_1(t_1) \wedge E_2(t)) \wedge |t - t_1| \leq T$$

A further issue to address involves the management of multiple occurrences of the same primitive event, during the detection of the composite event to which they belong. According to predetermined policies, it is possible to state which occurrences need to be considered during the composite event detection. Such policies, named *parameter contexts*, are used to set a specific consumption mode of these occurrences (collected in the Event History database). Four parameter contexts are defined in the Snoop event algebra. Given the concepts of *initiator* (the first constituent event whose occurrence starts the composite event detection) and *terminator* (the constituent event that is responsible for terminating the composite event detection), the four different contexts are described as follows:

- *Recent*: only the most recent occurrence of the initiator is considered;
- *Chronicle:* the (initiator, terminator) pair is unique. The oldest initiator is paired with the oldest terminator;
- *Continuous:* each initiator starts the detection of the event;
- *Cumulative*: all occurrences of primitive events are accumulated until the composite event is detected.

Therefore, the selection of specific parameter context states which component event occurrences play an active part in the detection process. Thus, the effect of the operators is conditioned by the context. The use of parameter contexts augments the semantics of the composite events and makes the detection mechanism very flexible with respect to different classes of applications (see section 3.3).



● Recent Context
○ Continuous Context
○ Cumulative Context
○ Chronicle Context

Figure 13 – Building of a composite event through graphical interface

DETECT is able to support the composition of complex events in the described EDL, through a *Scenario GUI* (Graphical User Interface). It is used to draw the event trees corresponding to threat scenarios, by means of an user-friendly interface (see Figure 13). Furthermore, the interface allows to specify:

- for each leaf node, the attributes “Event Identifier” and “Sensor Identifier”;
- for each operator node, the attributes “Temporal Constraint” (optional), “Alarm Level” (optional, see section 3.3 for further details), and “Any Parameter” (i.e. the *m* parameter required for the ANY composite event only).

The GUI is enabled to set further attributes, whose use is described in Chapter 5.

3.3 Composite event detection

The information sources in modern surveillance systems, in particular in extended infrastructures, may produce a very large number of events (warnings, alarms, diagnostic signals, and so on). In situation management task, this aspect can have a negative impact on:

- the capability to follow a stream of incoming events;

- the correct interpretation of events;
- the correct evaluation of seriousness and priority of events.

Composite event detection is crucial to recognize complex event patterns within the information flow. The formalism to represent composite event detection is graph-based, according to the Snoop event algebra [17][18]. In the described implementation, each graph is reduced to a tree, i.e. a directed acyclic graph with one root node. Therefore, the detection algorithm is able to reason on such data structure. Leaf nodes of the event tree represent primitive events. Internal nodes represent the composite events associated to the operators described in section 3.2. Each operator node can be considered as the root node of a corresponding sub-tree.

The event occurrence (both primitive and composite) flows bottom-up from the node to their parents. More in detail, primitive event occurrences enter the bottom nodes and flow upwards through the tree, being joined into composite event occurrences. In other words, for each primitive event occurrence, if its “Event Identifier” and “Sensor Identifier” are the same of the ones specified in the leaf node of the tree, the occurrence is propagated upwards. Going on with this approach, when the composite event – representing an internal operator node – occurs, it is propagated upward to the its parents, and so on. Therefore, the overall composite event (defined by the user) occurs when the process reaches the composite event associated to the root node of the tree.

As described in the previous sections, if a composite event has a temporal constraint, then it is propagated upwards only if the time interval between its component event occurrences fulfill that constraint.

The introduction of parameter contexts adds another perspective to the detection of composite events and solve the problem of the management of multiple occurrences of the same primitive event. The selection of this parameter states which component event occurrences play an active part in the detection process. The use of parameter contexts (Recent, Chronicle, Continuous, Cumulative) makes the detection mechanism versatile and flexible, with respect to different classes of applications:

- when the events happen at a fast rate and multiple occurrences of the same type of event only refine the previous detection;
- when there is a causal dependency between different types of events and their occurrences;
- when composite event detection along a moving time window needs to be supported;
- when all occurrences of constituent events are meaningful up to the occurrence of a deadline event.

According to the features of the application and to the objectives to fulfill, it is possible to select a consumption mode tailored to the needs. Once the occurrences are selected and used in the process of composite event detection, they are managed according to the following policies:

- recent: when the composite event is detected, all the component occurrences that cannot be initiators of that event in the future are flushed;
- chronicle: once used, all the occurrences of the constituent events cannot participate in any other occurrences of the composite event;
- continuous: a terminator event occurrence can cause the detection of one or more occurrences of the same composite event;
- cumulative: once used, all the occurrences of constituent events are flushed.

The main difference between the chronicle and the continuous contexts is that, in the former, for each initiator event there is a single terminator event, while in the latter multiple initiators can be paired with a single terminator [18]. To better understand the effect of each parameter context on the same composite event detection, it is possible to show the detection process on a timeline.

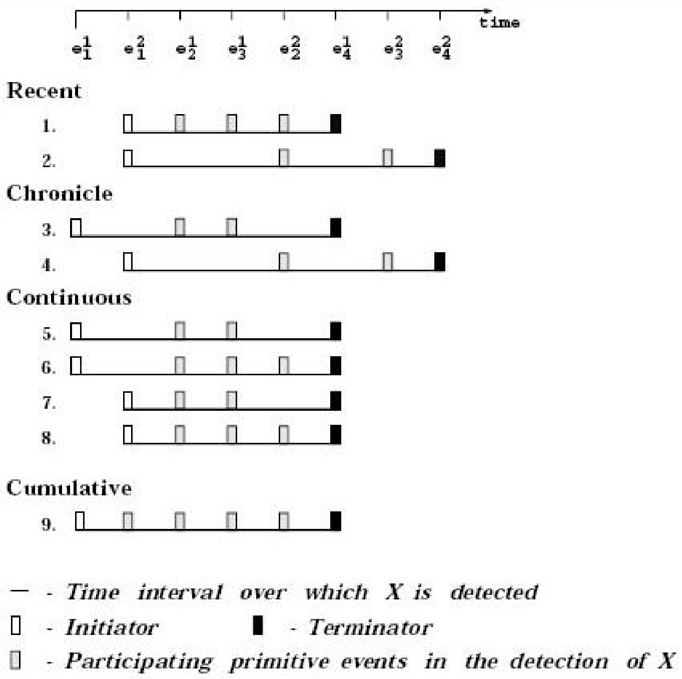

time
e_1^1 e_1^2 e_2^1 e_3^1 e_2^2 e_4^1 e_3^2 e_4^2
Recent
1.
2.
Chronicle
3.
4.
Continuous
5.
6.
7.
8.
Cumulative
9.
- Time interval over which X is detected
- Initiator
- Terminator
- Participating primitive events in the detection of X

Figure 14 – Composite event detection in different contexts

Given a chronology of primitive event occurrences, the Figure 14 shows an example of detection of the composite event X = (((E_1 AND E_2) SEQ E_3) SEQ(E_2 AND E_4)). The occurrences of each E_n event are denoted by e_n^c , where:

- *n* is associated to the event type;
- *c* is associated to the occurrence number.

The figure highlights the pairing mechanism between initiation and terminator events and how different instances of the same composite event are detected, given the sequence of primitive event occurrences and a specific parameter context.

An additional feature of the detection mechanism is the management of alarm levels, which the user can associate to a specific operator node, i.e. to a sub-tree, if he/she wants to detect the occurrence of such sub-tree. As mentioned in section 3.2, DETECT allows to associate an alarm level (different from 0) to each composite event, which should be signaled by the detection engine. In this way the user can be aware of the threat scenarios since their first evolution steps. Through such feature, the matching with known event trees to be recognized (stored in the Scenario Repository) can be also partial. Therefore, DETECT is able to detect whole threat scenarios and/or their parts, significant for the end-user, which require the early adoption of countermeasures, according to the evolution phase of the threat. Obviously, the type of countermeasure corresponds to the a-priori defined alarm level.

In the operational phase, when a composite event is recognized, the output of DETECT consists of:

- the identifier(s) of the detected/suspected scenario(s)[4];
- the temporal value related to the occurrence of the composite event (corresponding to the occurrence time of the terminator, given by the sensor timestamp);
- an alarm level (optional), associated to scenario evolution (used as a progress indicator and set by the user at design time);
- the identifiers of the primitive event occurrences, which led to the detection in object;
- other information depending on the detection model and/or on the sensors involved in threat detection (e.g. 'likelihood' or 'distance'). For further details see Chapter 4 and Chapter 5.

3.4 General architecture

This section presents the general architecture of DETECT, describing its component modules, their functionalities and their connections. The detection mechanism is mainly graph-based, according to the Snoop event algebra [17][18]. More specifically, in the described implementation, each graph is reduced to a tree. Therefore, the basic working logic follows the detection mechanism described in section 3.3, operating on

[4] The difference between detected and suspected scenario depends on the partial or total matching between the real-timeevent tree and the stored threat pattern.

the Event Trees which are composed with the EDL described in section 3.2. However, in order to achieve a more general architecture, within the design of DETECT, the representation of threat scenarios and the translation into detection models are separated from their resolution algorithms. So, the system is predisposed to manage different detection models and, for each of them, possibly different resolution mechanisms (if applicable). This aspect enhances in a significant way the flexibility of the system for future developments.

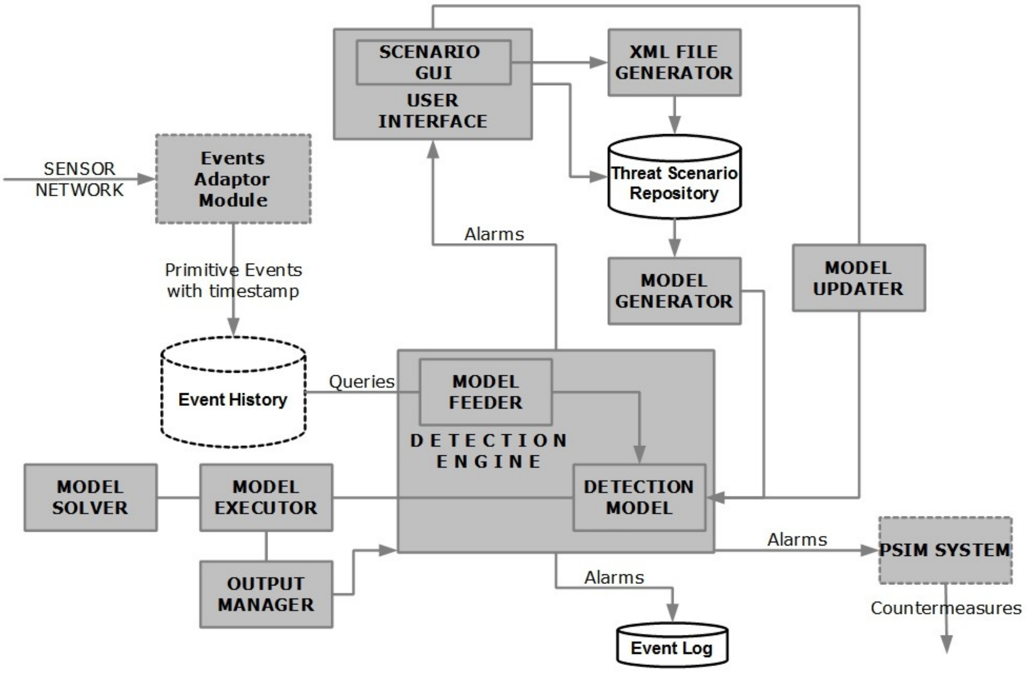


Figure 15 – DETECT architecture

The framework is made up by the following main modules (see Figure 15):

- **Event History**, that is database containing the list of occurrences of basic events detected by sensors, tagged with a set of relevant attributes including detection time, event type, sensor id, sensor type, sensor group, etc. (some of which can be optional). Since external sources (like a PSIM system) may populate the database, or a software bus may "virtually" represents the chronology of the Event History, it is depicted with dotted line. Anyway the minimal set of attributes which should characterize each occurrence of primitive events (in such a way to perform the correlation as described in the previous sections) has already been specified in Table 1. If necessary, a specific external **Events Adaptor Module** could aim to fill, in the right format, the Event History with the occurrences coming from the sensor network on the field.

- **Detection Engine**, supporting both deterministic (e.g. *Event Trees*) and heuristic models, sharing the primary requirement of real-time solvability. For each **Detection Model** there is a **Model Feeder** which instantiates the inputs of the engine according to the nature of the models by performing proper queries and data filtering on the Event History (e.g. selecting sensor typologies and zones). At the moment, the detection engine is only based on the deterministic model of the event trees, which are automatically fed whenever a new event occurrence is in the Event History.
- **Model Solver**, that is the existing or specifically developed tool used to execute the model. It implements the logical assumptions to solve the Detection Model, based on the inputs coming from the Model Feeder, therefore it is the responsible module for the composite event detection. We have implemented our own Model Solver based on the event trees Detection Model.
- **Model Executor** (one for each model), which triggers the execution of the mode, once it has been instantiated, by activating the related solver. An execution is usually needed at each new event detection.
- **Model Updater** (one for each model), which is used for on-line modification of the model (e.g. update of a threshold parameter), without regenerating the whole model (whenever supported by the modeling formalism).
- **Output Manager** (single), which stores the output of the model(s) and/or passes it to the interface modules.

- **Scenario GUI (Graphical User Interface)**, used to draw threat scenarios using an intuitive formalism and a user-friendly interface. Once a scenario has been built, it will be converted in a XML document by the **XML File Generator** module and then indexed in the **Threat Scenario Repository**. In this way we are able to store permanently all scenario features in a formal way as well as to facilitate possible subsequent data processing by other applications. In the opposite way, when the user selects the threat scenario he/she wants to detect from the repository, the XML document which contains its description has to be re-converted in the detection model, which represents the composite event related to the scenario. This task is carried out by the **Model Generator** which recovers the original graph and its parameter by parsing the related XML document.
- **Event Log**, which is kept to gather all information about detected events (detection time, alarm level, instances of component events involved in the composite event detection process, and further information like "alarm reliability" and "distance" with respect to other items of Threat Scenario Repository, see Chapter 4 and Chapter 5). Detected alarms could be also sent to existing PSIM systems in order to trigger adequate countermeasures.

3.5 Advantages and limitations

DETECT can be used as an on-line decision support system, by alerting in advance PSIM system operators about the likelihood and nature of the threat, as well as an autonomous reasoning engine, able to guide the activation of response actions. The latter include, for example, audio and visual alarms, emergency calls to first responders, air conditioning flow inversion, activation of sprinkles, etc. The DETECT architecture is inherently suited to many application domains: not only CIP and HS, but more in general all the fields like environmental monitoring and control [34]. The framework is being experimented in railway transportation systems, which have been demonstrated by the recent terrorist strikes to be among the most attractive and vulnerable targets. Real threat scenarios include intrusion and drop of explosive in subway tunnels, spread of chemical or radiological material in underground stations, combined attacks with simultaneous multiple train halting and railway bridge bombing, etc. DETECT has proven to be particularly suited for the detection of such articulated scenarios, using a modern PSIM platform, in turn based on an extended network of cameras and sensing devices (see Chapter 7).

The use of a simple formalism (apparently not powerful, with respect to other), based on Event Tree models, makes light, efficient and easy-to-use the whole approach. Such features are crucial to assure the usability of DETECT and to satisfy the requirements described in section 2.1 (like the real-time solvability of the correlation models). At the same time, the tool doesn't ask for a modeling expert to draw event trees, given their intuitiveness. Accordingly, this helps in making the models easy to update and integrate by the security operators, reducing and simplifying the maintenance effort as well. The same considerations are not valid for more powerful formalisms (e.g. ANN, Petri Net), which are far from being straightforward to implement, control and update. However, the general architecture of the framework is suitable to accommodate different detection models, which could be used in parallel with the event trees. With respect to traditional approaches of infrastructure surveillance, DETECT allows for:

- A quick and focused response to emergencies, which could be fully automatic or dependent on human supervision and intervention. A semi-automatic approach may represent the right trade-off, since human management of critical situations, possibly involving many simultaneous events, is a very delicate task. Furthermore, it can be error prone as well as subject to forced inhibition.

- An early warning of complex threat scenarios since their first evolution steps using the knowledge base provided by experts during the qualitative risk analysis process. This allows for preventive reactions which are very unlikely to be performed by human operators given the limitation both in their knowledge base and vigilance level. Therefore, a greater situational awareness can be achieved.
- An increase in the Probability Of Detection (POD) while minimizing the False Alarm Rate (FAR), due to the possibility of logic as well as temporal correlation of events. While some PSIM software offer basic forms of logical correlation of alarms, the temporal correlation is not implemented in any nowadays systems, to the best of our knowledge (though some vendors provide basic options of on-site configurable "sequence" correlation embedded in their multi-technology sensors).

On the other hand, the main limitations of the framework are inherently linked to the introduced deterministic approach, where completeness and correctness of the knowledge base depend on the quality of the Risk Analysis. Therefore, the search for an exact matching with the items of the knowledge base should be extended (alarm level management is not enough). In addition to that, there is a lack of awareness about alarm credibility and likelihood, that is crucial to understand priority of intervention and react consequently. A probabilistic approach should complement the deterministic one to manage also the uncertainty coming from all the information sources (single sensors, multimodal monitoring systems, integrated surveillance systems, as well as detection and correlation models). Chapter 4 and Chapter 5 describe how to improve the detection effectiveness and efficiency.

Chapter 4

Heuristic detection of threat scenarios with Event Tree distance metrics

4.1 Problem statement

The following section introduces an extension of the described event correlation approach with event tree similarity analysis capabilities. It enables an earlier and more robust recognition of threat scenarios, due to the possibility of detecting sequences of events with a non perfect matching, as well as an increased tolerance to sensor and modeling faults.

In the context of situation recognition, a common technique to address that issue is the graph matching. It usually based on the extraction of a relational structure from the graph and on the comparison of the extracted structure with a set of stored structures of interest for the application to find the best match [75]. However, graph matching problems are typically NP-complete, thus the related algorithms have problems with performance [76]. In fact, in the worst case, the time required to execute the algorithms increases exponentially with the size of the graphs. Therefore, this aspect reduces the field of application of many techniques based on graph matching. Algorithms with lower complexity have been studied, but they often introduce several constraints to fulfill (e.g. topological restrictions to the compared graphs). In addition to the complexity, further problems include the difficulty in representing and recognizing situations and relations, which may also have temporal constraints. Therefore, graph matching techniques are often used only for forensic applications and post-event analysis.

The techniques based on similarity between graphs, representing threats and attacks, are not new in particular in the field of computer network, where a similar issue to the one addressed in this section is to recognize threats even in the case of missing events. A widespread solution in network intrusion detection systems is the following. If an alert, which is the known consequence of a forerunning event, is received by the correlation engine and if the forerunning event has not been detected, then the missed event can be identified. However such a solution has an intrinsic limit: it cannot cope with missing events that are not linked to other events.

The approach described in the following differs from all these techniques and it is consistent with the indications of section 2.1. It does not require: a) to satisfy constraints on size or topology of the event trees, b) direct cause/consequence relationships between detected and missing events, since it is based on ad-hoc metrics. The latter allow for a heuristic recognition of similarities between event trees. The analysis can be performed both at the insertion time of a new detection model into the engine (off-line mode), and at the run time of the engine (on-line mode). To the best of our knowledge, no existing physical security monitoring system features a scenario-based heuristic detection approach, like the one described in this chapter. With respect to traditional approaches of infrastructure surveillance, the framework enriched by this extension allows for a more robust and straightforward early warning of complex threat scenarios, and a more rational use and management of the knowledge base provided by risk analysts.

4.2 Definition of distance metrics

The approach requires the definition of ad-hoc metrics distances. As stated in Chapter 3, each event tree consists of basic events (the ones detectable by each sensorial subsystem) and the connectors used to associate them (to express logic, spatial or temporal relationships). The former are the leave nodes, the latter are the internal nodes of the event tree. Furthermore, it is possible to specify additional attributes, related to the whole tree (e.g. the parameter context) or its nodes (e.g. to set the type of connector, temporal constraints, etc.). That is the reason why a complete comparison between event trees should involve, in addition to the structure (to find a possible isomorphism), also the above mentioned attributes. Therefore, it is necessary to define appropriate metrics to evaluate the distance between event trees, in terms of event trees structure (e.g. number of nodes), attributes of nodes (e.g. event type for leaves nodes; connector type for internal nodes) and attributes of trees (i.e. the parameter contexts). More formally, the following attributes can be associated to Event Trees (in the form of positive integer numbers):

1. *TN*: total number of nodes
2. *TD*: tree depth, that is the number of levels from leaves to the top node
3. *TW*: tree width, that is the maximum number of operators at the same level
4. *SL*: set of leaf nodes
5. *SO*: set of operator nodes

Though other attributes (e.g. number of arcs) could be associated to event trees, the ones listed above picture a comprehensive yet not redundant set of characteristics. While a theoretical demonstration could be possible, such a statement has been validated experimentally. For instance, the number of arcs in all the significant scenarios included in the repository was always dependant on the number of nodes.

In order to obtain an easy to compute metric, the distance between two event trees is obtained as the sum of the differences between homologous attributes. In other words, the distance *D* among event trees *A* and *B* is obtained as follows:

$$D = |TN_A - TN_B| + |TD_A - TD_B| + |TW_A - TW_B| + DSL_{AB} + DSO_{AB}$$

(+ 1 if parameter contexts are different)

The quantities DSL and DSO are computed as set differences (*card* competes the cardinality of the set):

$$DSL_{AB} = card\,(SL_A \cup SL_B) - card\,(SL_A \cap SL_B)$$

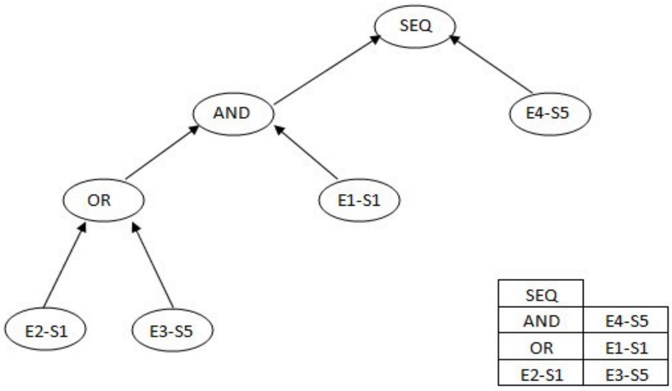
$$DSO_{AB} = card\,(SO_A \cup SO_B) - card\,(SO_A \cap SO_B)$$

It is quite obvious that such a heuristic distance metric can be applied to any couple of event trees, regardless of possible isomorphism[5].

4.3 Implementation in DETECT

In order to compute tree attributes, an appropriate algorithm has been implemented in DETECT. Starting from the root node, the whole tree is scanned and each node is saved in a table where each row represents a tree level (see Figure 16). For each node, the name of the father node is saved as well as the list including the names of the son nodes. In the end of the scan, all the information relevant for tree attributes computation will be available. Hence the formula to obtain the distance between any couple of trees can be easily computed. Off-line distance calculation is very useful when inserting a new event tree in the Scenario Repository. In fact, when a human operator finishes building the event tree and saves it in the repository, he/she can see all the distances (possibly only the ones lower than a certain threshold) with all the other event trees in the repository. Therefore, if another tree exists whose distance from the new one is very low, then it is possible the two trees represent the same threat (or similar threats) and therefore could be somehow merged to reduce multiple warnings and improve usability as well.

[5] Two trees are isomorphic when they are identical in graph structure (they could differ in node attributes).



SEQ	
AND	E4-S5
OR	E1-S1
E2-S1	E3-S5

Figure 16 – An example of table obtained from an event tree

For on-line calculation, the distance needs to be computed bottom-up starting from subtree attributes, which will be associated at run-time to each node (see Figure 17). Due to the working logics of DETECT, some limitations hold for the run-time computation of tree attributes (e.g. the *TD* metric cannot be computed at run-time). More specifically, since operator nodes can be considered as the roots of the subtrees below them, it is possible to associate to operator nodes the attributes of the subtrees below them. Hence, moving from the leaves to the root and exploiting the already computed attributes, each operator node will be associated to updated attributes representing all the tree structure below it. Therefore, the root node will include the overall attributes of the whole tree. When a subtree is detected and its alarm level in DETECT is greater than 0, its attributes are compared with the ones of all the other event trees in the Scenario Repository. If the distance D with another threat scenario T is lower than a configurable threshold D_T , then a warning is generated and shown to the PSIM human operator, in order to warn him/her about the risk that threat T is occurring. It is obviously possible to associate different warning levels to different distances (the lower the distance metric, the highest the warning level); however, in practical applications it is important to keep the system simple to understand to operators. Therefore, we have decided to use a single threshold and a single warning level. An application example of the heuristic approach is introduced in Chapter 6.

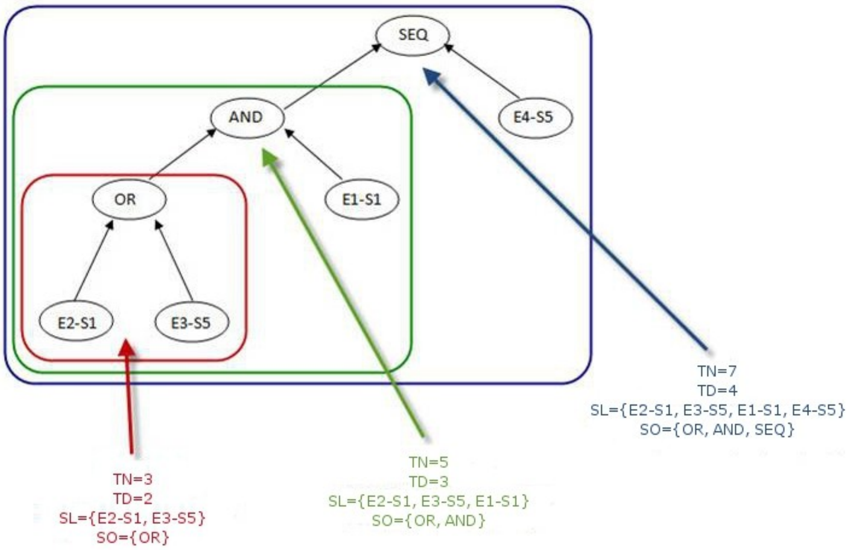


Figure 17 – Example of on-line subtree attributes computation

4.4 Benefits and practical implications

The described technique allows to achieve several important results. In off-line mode, the approach allows to provide an effective feedback to the experts responsible of the identification of threat scenarios, in the phase of Risk Analysis. In such a way, they are able to recognize and study possible classes of equality in the identified scenarios, not arisen before. Furthermore, DETECT allows to store in the Scenario Repository only the event trees actually corresponding to new patterns, since it detects possible redundancies when updating the Scenario Repository. The consequence is a global improvement of the knowledge base on which the correlation engine works.

The same approach is applicable also during the phase of post-event analysis to support security operators. In fact, let us assume that a specific attack has already occurred in the monitored environment, but the integrated framework has not recognized it. That is probably due to a lack in the knowledge base. In other words, the corresponding event tree representing the occurred threat scenario was not yet in Scenario Repository. If we know some information about the possible dynamic of the occurred scenario, we can also model it, define the related tree, and activate the correlation engine on the temporal window of interest (e.g. to detect suspicious behaviors and, consequently, identify the possible attackers). For example, we may be interested in quickly selecting all the multimedia streams of interest. Such a feature enhances the post-event forensic search of traces of an attack scenario not previously stored in the Scenario Repository. The distance between the just defined tree and the already stored ones can provide useful indications regarding this kind of analysis.

In on-line mode, DETECT has the responsibility of performing queries on the Event History for the real-time feeding of detection models and of recognizing the complex events stored in the Threat Scenario Repository. The ordinary working logic consists of verifying the matching between trees, step-by-step, each time a new basic event is taken from the Event History. Depending on the partial or total matching between event trees, the framework is able to report warning messages related to detected/suspected threat scenarios. In addition to that, exploiting the described approach, it is possible to extend the recognition capabilities of the framework. In fact, on-line heuristic detection also evaluates the similarity of the partial trees constructed on-the-fly (while the correlation engine runs) with the known event trees, stored in the repository. In this way, the system is able to show information regarding possible distances with respect to known event trees and to give useful warnings about possible threat scenarios to human operators. The existing "deterministic" approach is hence extended and the recognition is more robust to both imperfect scenario modeling, due to limitations in the human knowledge, as well as to possible missed detections by sensors, which are not 100% reliable.

The further advantages of similarity matching lie in the inner early warning capability, not requiring further modeling efforts, since scenario matching is not required to be complete, nor exact. The operators can then evaluate, through the user interface, the warning level of suspected threats, which is inversely proportional to the computed distances. The overall effect achieved is a higher level of security since the quantitative indications about unfolding scenarios allow operators to quickly undertake appropriate countermeasures.

Chapter 5

Probabilistic evaluation of detection trustworthiness

5.1 Problem statement

Modern surveillance solutions for infrastructure protection are based on the integration of different sensing subsystems. Each subsystem can include a large number of diverse and/or redundant distributed sensors, which are in charge of detecting abnormal conditions or unwanted events in the monitored environment. The rational exploitation of the available sensing capabilities needs a proper management and processing of both the "modeled" and "captured" information together with the related uncertainty. Therefore, together with PSIM systems there is an increasing need for the appropriate management of parameters characterizing sensors performance [34] [37].

Ideally, the sensors should detect only "real" alarms, that represent a true threat. However, many devices generate unnecessary warnings, which can be classified as false alarms or nuisance alarms. False alarms are due to events that should not cause an alarm, while nuisance alarms are generated when a legitimate cause occurs, but alarm activation is not due to a real threat. The same consideration is still valid for a sensing subsystem as a whole, i.e. including sensing devices and specific software for the processing of what they detect (e.g. intelligent video surveillance systems include cameras and video content analytics for the detection of events).

The aim of this chapter is to describe the means to improve efficiency of situation recognition provided by DETECT, which are generically applicable in the PSIM context. Efficiency is to be intended at human-machine interaction level, by associating a level of trustworthiness to threat detection in order to allow human operators to be aware of alarm credibility and priority of intervention, and hence react consequently.

More in detail, we can evaluate the impact of the reliability of each sensor/subsystem on the reliability of the whole integrated surveillance system, in terms of POD (Probability of Detection) and FAR (False Alarm Rate) parameters. The first characterizes the effectiveness of a detection system, the second determines its operational viability [82][83]. The need for such an evaluation is especially important when integrated surveillance systems are extended by means of a correlation engine aimed at the automatic threat detection and situation recognition. In fact, in that case, the alarm activation is based on the correlation of different sensors output. Furthermore, the alarms can be sent to a control center and can involve the triggering of specific countermeasures, by means of a fully automatic (independent from human intervention) or partially automatic (based on human discrimination) procedures.

In order to fulfill such an objective, the deterministic approach described in the previous chapters should be complemented by probabilistic ones. The latter are very popular in the scientific research, in particular to recognize situations as well as threats in uncertain environments. More in detail, the approach is based on the application of Bayes' theorem, used to evaluate the degree of belief of a hypothesis H (i.e. an alarm activation), given observed data E (i.e. the event triggering the corresponding alarm has been observed). Therefore, the aim is not to perform a new inference, with respect to the deterministic approach, finding the most likely hypothesis between all the possible ones that may explain the empirical evidence. Instead, it consists of evaluating the trustworthiness of the inference provided by the deterministic approach, taking into account the uncertainty due to the sensors. In order to represent composite events, the formalism is based on a Bayes Network (BN).

5.2 Uncertainty in threat detection

This section introduces an additional feature to quantify the uncertainty due to sensor false alarms. In particular it focuses on how to exploit the parameters describing the detection performance[6] of the sensors, which are involved in physical security situation recognition. As mentioned above, the aim is to evaluate the trustworthiness of the inferred alarms.

In order to associate a reliability level to event detection, it is possible to use a real-time fuzzy correlation of sensor outputs using a Bayesian Network (BN). Such a probabilistic modeling formalism enables a fuzzy logic through the use of "noisy" logic gates, whenever the output is not deterministic, but associated with a certain probability [56].

Formally, let us define a *detector* as a sensor or a sensing subsystem which in relation to a certain event can provide two outputs:

- TRUE – if the event has been detected;
- FALSE – if not.

Each detector can be associated to the following parameters:

- POD = P(event detected | event occurred);
- FAR = P(event detected | event not occurred).

An analysis based on the POD of detectors can be used to compute the probability in threat recognition, while we build the related detection models. Therefore it is convenient at design-time, since the results can provide a guide to draw appropriate event trees and to support the choice and dislocation of detectors, with respect to the specific threats to be addressed. The main end is to reach a certain target in the probability of recognition a particular threat, before using its detection model at real time. Such an analysis is objective of another work and it is not described in this thesis. Let us to address a FAR based real-time analysis in the following.

[6] In this section we refer to detection performance, reliability and trustworthiness by meaning the same concept related to false alarm generation (i.e. false positive).

Assuming the use of AND logical operator in order to correlate the outputs of detectors, we can perform an analysis based on their FAR parameters and aimed at the calculation of alarms reliability in real-time. A synthetic indicator of such an evaluation can then be reported to human operators together with inferred alarms. To better understand the approach we proceed with an explanatory example.

Detector ID	Event ID	FAR
S1	E1	0.15
S2	E2	0.10

Table 2 – Probabilistic parameters of two possible sensors

In the following, we assume using two detectors whose FAR is described in Table 2. With reference to the AND operator, we can model the alarm reliability through a simple Bayesian Network (see Figure 18).

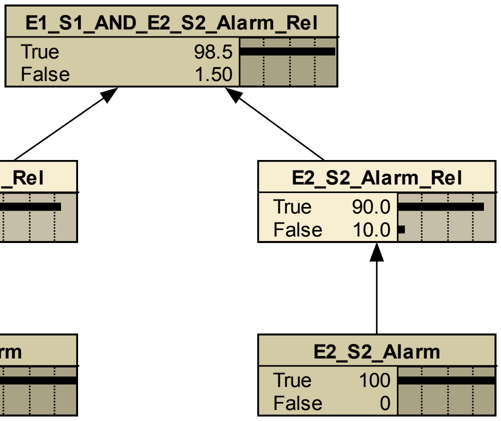


Figure 18 – Example of BN modelling an AND logical operator

The leaf nodes represent the occurrence of the alarms associated to the events E*x* detected by S*x*. The reliability of each alarm (E*x*_S*x*_Alarm_Rel) is calculated using the FAR parameter of the related detector. The corresponding formula is the following:

P(Ex_Sx is TRUE | Ex_Sx_Alarm has been generated) =

= P(Ex_Sx_Alarm is not FALSE) = 1 − P(Ex_Sx_Alarm is FALSE) = 1 – $\text{FAR}_{\text{E}x\_\text{S}x}$

The alarms reliability reported in the BN are represented in percentages. The CPT (Conditional Probability Table) of the AND node is reported in Table 3.

E1_S1_Alarm_Rel	E2_S2_Alarm_Rel	E1_S1_AND_E2_S2_Alarm_Rel
True	True	True
True	False	True
False	True	True
False	False	False

Table 3 – CPT of the AND node

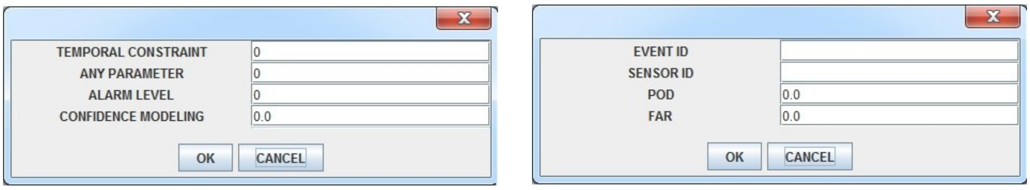
The following hypothesis holds: the alarm associated to the AND event is not considered reliable only if both the alarms associated to E1_S1 and E2_S2 events are not considered reliable. For example, it means that when S1 detects E1 correctly and S2 generates a false alarm in E2 detection, then the related AND event – which is triggered anyway – is classified as TRUE. However, by modifying the CPT properly we can consider a more conservative hypothesis: the alarm associated to the AND event is considered TRUE only if both the alarms associated to E1_S1 and E2_S2 events are considered reliable. In the first case (shown in Figure 18) we have an AND alarm reliability of 98.5%, in the second one, we have a lower value (76.5%). Therefore, according to the protection strategy to be pursued, we can set the CPT of the AND node.

5.3 Implementation in DETECT

The theoretical discussion in the previous section is suitable for a simple application to the composite event detection described in Chapter 3. To evaluate in real-time the trustworthiness of detected alarms, which correspond to composite events, the main issue is to reason on the applicability of the approach to the EDL operators (see section 3.2). In particular, understanding how to use the BN in all the possible cases is needed. On the contrary, once the composite events are detected, temporal constraints and parameter contexts of the EDL have no impact on this type of analysis.

The real-time calculation of an OR alarm reliability is quite simple. In fact, OR alarm activation is concomitant with the single E*x*_S*x* alarm generated first. Considering the example of the previous section, the result is 85% or 90% depending on the case. The approach is easy to apply also to the other operators. In fact, in real-time analysis, the SEQ (sequence) operator can be treated as an AND. The sequence operator considers the temporal order of the constituent primitive events, however (once SEQ is occurred) that aspect is not significant anymore. From the viewpoint of such evaluation, we always have to consider that trustworthiness depends on the reliability of the two sensors, as it happens in the case of AND operator. It is quite obvious that the same way of reasoning can be extended to the case of ANY operator. In fact, it occurs when *m* out of *n* distinct events specified in its expression occur, regardless the temporal sequence. Therefore $ANY(m, E_1, E_2, \ldots, E_n)$, which is equivalent to the "m out of n" scheme, can be treated (once occurred) as an *n*-ary AND. Whatever the set of *m* distinct occurred events, the reliability of the *m* sensors that have detected them is the only significant item for the evaluation of ANY detection reliability. Accordingly, the BN associated in real-time to the ANY activation will be again the same, but with *m* branches.

The effectiveness of the approach increases significantly when we consider more complex Event Trees. In those scenarios, when primitive events are detected by sensors, they feed detection models according to the scenario evolution. Thus, step by step, the BN related to each occurred subtree can be executed in real-time in order to get also the alarm reliability related to the inferred composite event.



TEMPORAL CONSTRAINT: 0
ANY PARAMETER: 0
ALARM LEVEL: 0
CONFIDENCE MODELING: 0.0

EVENT ID:
SENSOR ID:
POD: 0.0
FAR: 0.0

Figure 19 – DETECT entry windows for operator and basic event parameters

Finally, we can take into account also the uncertainty of the detection models used to recognize threat scenarios. More in detail, in order to consider a possible mismatching between a real threat scenario and its model, for each logical operator there is the possibility to set also a confidence index (named *confidence modeling*), which weighs the trustworthiness of the operator. In other words, at the design time of the detection model (i.e. the event tree and all the related parameters) the responsible can set the index of certain EDL operator to a probability value p in the range from 0 to 1 (1 is the default value representing no uncertainty). Hence, the occurrence of the logical condition represented by the EDL operator, will be True with a probability p weighted with the computed alarm reliability.

All the input parameters of the nodes, introduced up to this point, can be entered in DETECT framework by means of proper windows of its GUI (see Figure 19). The BN computation for a composite event occurrence is performed only if it corresponds to a detected or suspected threat scenario[7], and the FAR parameters of all the involved detectors are available (whereas the default value of confidence modeling of EDL operators is 1, if not specified by the user). Furthermore, the computation is fully automatic, embedded in the framework and doesn't require extra effort from the user. In the current implementation, DETECT takes into account only the less conservative version of the CPTs of the EDL operator nodes. The possibility to select one of the two modes (like described in the previous section) will be considered in the next developments.

[7] In case of occurrence of whole event trees or sub-trees with an alarm level different from 0.

5.4 Benefits and practical implications

As stated in section 3.5, the main advantages of DETECT consist of an early warning of complex threat scenarios as well as an enhanced detection reliability, which include in particular the lowering of the False Alarm Rate in composite event detection. The latter is a direct consequence of logic, spatial and temporal correlation of events, detected by redundant and/or heterogeneous devices. Thanks to the described contribution is also possible to provide evidence of that, and to quantify the achieved "gain" (e.g. section 5.2 shows that AND correlation has a resulting FAR equal to 1,5%, with respect to 10% and 15% of the single sensors).

Furthermore, the trustworthiness evaluation is strategic in the interaction with a human operator and in the activation of countermeasures. Recent studies in PSIM context remark the importance of having tools that can identify, through a stream of device data, situations in real-time, as well as sort and prioritize them [79]. Therefore, the awareness of alarms credibility supports the decision on the priority of interventions. From this point of view, it is a delicate task for two reasons. From one side, we may have a fully or partial automated operating procedure, in order to react to the recognized situations. Since one or more steps are not confirmed by human operators, a basic requirement is to undertake a countermeasure only when there is the reasonable certainty that it is necessary.

On the other side, whatever is the automatism of the response procedures, their execution should always be proper. For example, let us consider different response actions to an occurred threat scenario: opening specific exit gates and closing specific entrance gates in a public area (e.g. for conveying a mass of people in a certain direction in case of emergency), sending security staff on the interested site, emergency calls to first responders, and so on. According to that, more critical is the countermeasure, more important is to assure that it is necessary. Thus, its activation could be confirmed only if the correspondent alarm reliability exceeds a certain threshold (e.g. specified by the end-user). That is a basic requirement to optimize the impact of the protection efforts and to manage all the available resources efficiently. The fulfillment of these needs can have important practical implications, in terms of limited costs for the security and overall organizational efficiency, which is fundamental to assure an adequate protection level.

Chapter 6

Applications in the railway and mass-transit domains

6.1 Modeling of a threat scenario

This section reports an example of application of the overall approach to a case-study in a metropolitan railway environment [33]. Historically, these mass transit systems, being easy to access public places, are vulnerable to many threats of various kind and seriousness. In fact, they can be theater of criminal acts, aggressions, vandalism as well as sabotages and terrorist strikes. The following is a description of how to detect complex scenarios of terrorist attacks by exploiting heterogeneous sensing devices.

Modern smart-surveillance systems suitable for the protection of metro railways are made up by several non fully reliable sensorial subsystems. When single alarms are not reliable, automatic countermeasures cannot be activated and operators response is slowed down. Mechanisms of alarm correlation can contribute to reduce the FAR (False Alarm Rate) and at the same time improve the POD (Probability of Detection). Improvements in detection reliability can be achieved adopting two main techniques: redundancy and diversity.

Through complex computer vision algorithms, the video analytics allows for the detection of events of different complexity, like intrusions in critical areas, abandoned objects [47], abnormal behaviors (person running or loitering, downfalls, etc) [50]. Since the detection of an event can suffer from the intrinsic reliability of the algorithm, as well as from issues due to environmental conditions (e.g. changes of lighting, presence of reflective materials, occlusions), **redundancy** in cameras dislocation can improve detection reliability and overall system resiliency against both accidental and intentional faults. For example, assuming the use of more intelligent cameras with overlapped views from different viewpoints to detect an abnormal behavior in a platform, the events detected by each camera can be combined with a simple AND logic.

However, the most interesting application of redundancy is when it is used in combination with **diversity**, by exploiting devices based on different technologies. In the assumption that the abnormal behavior includes screaming, which is detectable by means of appropriate audio sensors, the information coming from the microphone and the cameras installed in the platform can be combined using a more complex approach, based on the use of advanced logical and temporal operators.

Let us suppose to address a chemical attack, similar to what happened in the Tokyo subway on March 20, 1995 using Sarin gas. Sarin is a chemical warfare agent (CWA) classified as a nerve agent. It is a clear, colorless, odorless, and tasteless liquid in its pure form, and can evaporate and spread in the environment very quickly.

The current available technologies to identify the contaminated areas, for example include Ion Mobility Spectroscopy (IMS), Surface Acoustic Wave (SAW), Infrared Radiation (IR), etc. They are employed in ad-hoc standoff detectors and each of them is characterized by different performances. One of the most accurate device, the automatic scanning, passive, infrared sensor can recognize a vapor cloud from several kilometers with an 87% detection rate [25]. Thus, to improve sensitiveness and reduce the number of false alarms, different technologies are often integrated in the same standoff detector (for example, the IMS and SAW detection are typically combined). More in general, it is possible to combine heterogeneous detectors and to correlate their alarms (e.g. IMS/SAW and IR detectors), in such a way to get an early warning system for the detection of chemical agents. Exploiting the redundancy and diversity also of these devices, increasingly complex correlations (logic, temporal, and spatial) can be implemented.

A likely scenario consists of a simultaneous drop of CWAs in many subway platforms in the rush hour. Let us suppose that dynamic of events is the following:

1. the attackers stay on the platforms, waiting for the simultaneous drop of CWA;
2. the first contaminated people fall to the floor;
3. the people around the contaminated area run away and/or scream;
4. the CWA quickly spread in the platform level and reach the escalators to the concourse level.

In each subway site, it is possible to address the attack scenario by means of two intelligent cameras positioned at platform end walls, a microphone between them, two standoff detectors for CWAs positioned on the platform and on the escalator.

The scenario can be formally described by means of the notation “sensor description (sensor ID) :: event description (event ID)”:

Intelligent Camera (S1) :: Fall of person (E1)

Intelligent Camera (S1) :: Abnormal running (E2)

Intelligent Camera (S2) :: Fall of person (E1)

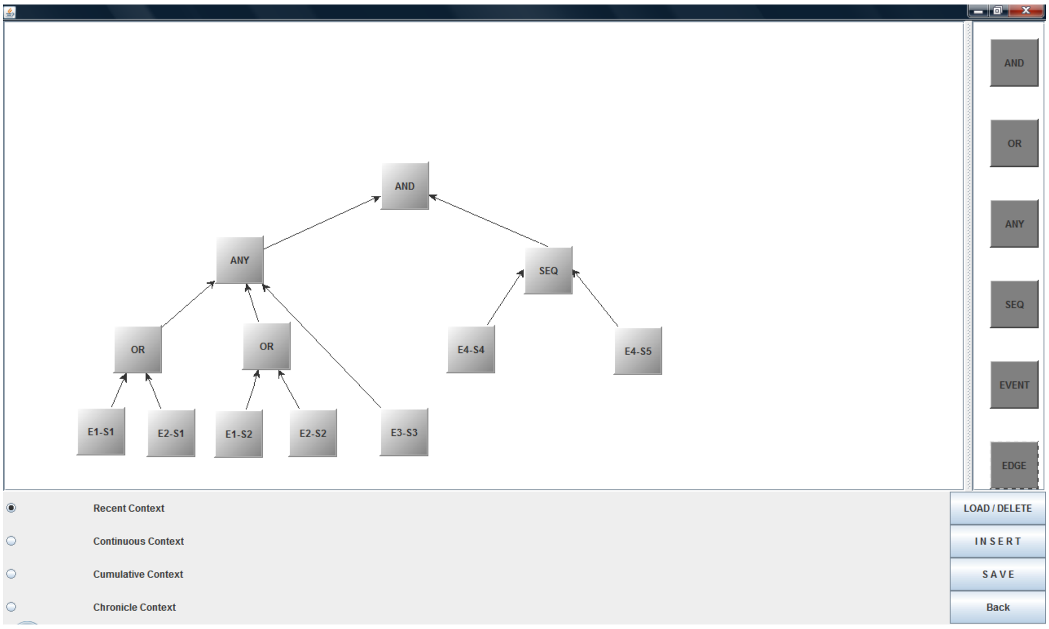
Intelligent Camera (S2) :: Abnormal running (E2)

Audio sensor (S3) :: Scream (E3)

IMS/SAW detector (S4) :: CWA detection (E4)

IR detector (S5) :: CWA detection (E4)

Given the scenario described above, the composite event **drop of CWA in platform** can be represented by the event tree in Figure 20, built using the DETECT framework.



● Recent Context

○ Continuous Context

○ Cumulative Context

○ Chronicle Context

Figure 20 – Event tree associated to "drop of CWA in platform" by using DETECT

Please note that single events detected by intelligent cameras do not represent necessarily a threat situation. In the approach we are describing, a low alarm level (e.g. to 1) can be associated to the OR operators.

A partial alarm level (e.g. 2) can be associated to the scenario evolution in case of occurrence of the ANY event (at the left of tree). The *m* parameter of ANY is set to 2 (trough the Scenario GUI) , this means that when 2 out of 3 distinct events detected by intelligent cameras and/or microphone occur, the monitored situation is considered abnormal (in fact each of the single events: person who falls, runs or scream, can not represent a meaningful state of alert). Besides, a temporal constraint can be set on ANY operator, in such a way to catch real alarm conditions: e.g. if fall and scream are detected at a distance of time of 30 minutes, that could not represent an alert condition for the specific scenario. In the specific example, it could be set to 5 minutes to take into account the latency of both gas propagation and intoxication symptoms.

An higher alarm level (e.g. 3) can be associated to the scenario evolution in case of occurrence of SEQ event (at the right of tree). The use of the sequence operator is due to the different assumed locations of the CWA detectors: IMS/SAW detector at platform level, IR at escalator or concourse level, in such a way to detect correctly the spread of CWA. If IR detector gives a warning before the one based on IMS/SAW detection, this could be an abnormal condition due to a false alarm and should not cause the activation of a warning. To further avoid false alarms, also a temporal constraint should be set. In this case, it can be set to 10 minutes to be conservative while taking into account the movement of air flows between different environments.

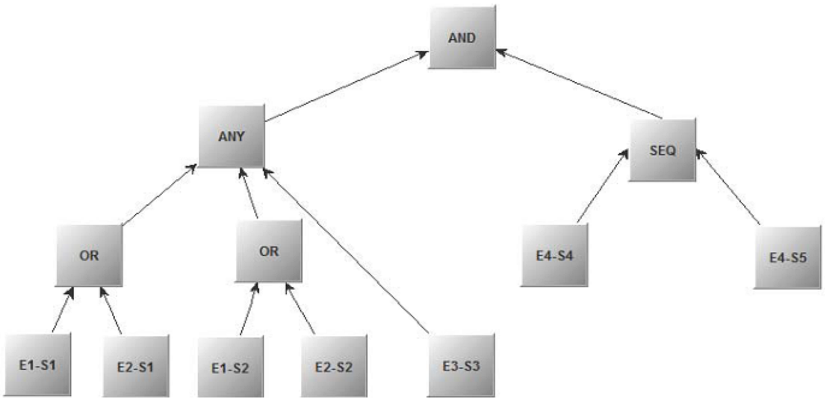
The detection of the whole threat scenario is associated to the AND occurrence. Its alarm level is set to 4. Finally, it is necessary to set the parameter context to regulate the consumption mode of the occurrences of events in feeding the detection engine. In this case, the assumption is that only the most recent occurrence of each event is meaningful. Thus, parameter is set to "recent context".

The use of many alarm levels is strategic to trigger countermeasures properly: e.g. the alarm level 2 can trigger the opening of the turnstiles; at level 3 an appropriate ventilation strategy can be activated; finally, the detection of the whole composite event can be associated to actions like: evacuation message from public address, stop trains from entering the station, and emergency call to first responders [31][33].

6.2 Distance metrics computation

In this section we report some examples of evaluation of attributes and distance metrics, as described in Chapter 4, for reference threat scenarios [31]. The first scenario we consider is the Chemical Attack (scenario A) by means of a CWA (Chemical Warfare Agent), which we have already described in section 6.1, whose event tree is depicted in Figure 21 together with a table including its attributes.

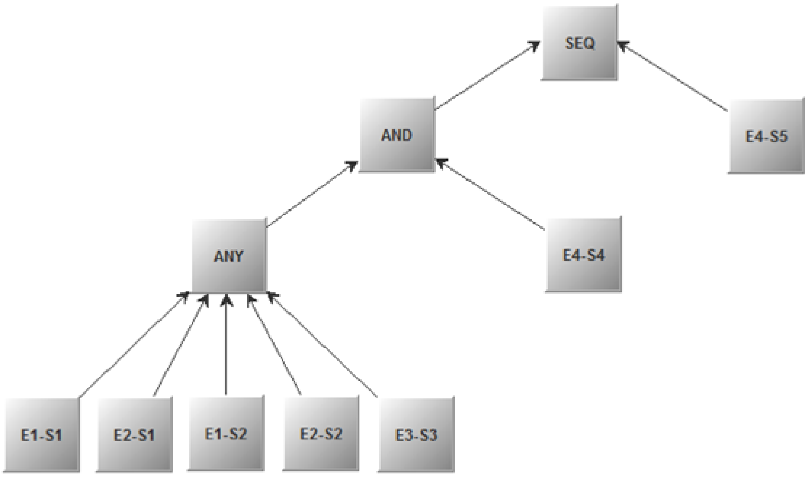
The same scenario could be represented in other way using the model of Figure 22 (scenario B), featuring slightly different attributes.



SCENARIO A

TN	12	
TD	3	
TW	2	
SL	E1-S1, E2-S1, E1-S2, E2-S2, E3-S3, E4-S4, E4-S5	cardinality=7
SO	AND, ANY, SEQ, OR	cardinality=4

Figure 21 – Event tree attributes for the Chemical Attack scenario



SCENARIO B

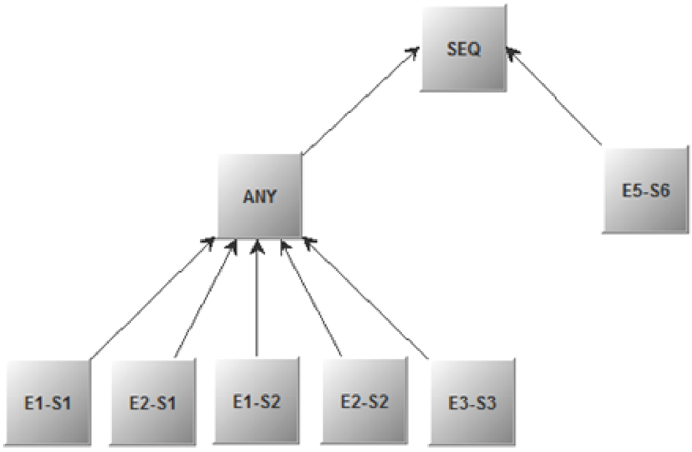
TN	10	
TD	3	
TW	1	
SL	E1-S1, E2-S1, E1-S2, E2-S2, E3-S3, E4-S4, E4-S5	cardinality=7
SO	SEQ, AND, ANY	cardinality=3

Figure 22 – Event tree attributes for another version of the Chemical Attack scenario

The two scenarios of Figure 21 and Figure 22 feature the same primitive events (i.e. the trees have the same leaves) and therefore the SL distance is 0. Instead, the sets of operators differ by 1. Overall, the distance is given by:

$D = |12-10| + |3-3| + |2-1| + 0 + 1 = 4$

Now, let us consider a scenario of pickpocketing/aggression (scenario C), which could partially overlap with the previous one regarding people behaviour, since it features the composite event represented by the ANY operator included in scenario B. Furthermore it is similar to the corresponding ANY in scenario A. Please refer to Figure 23, where E5-S6 represents an alarm coming from the emergency call point.



SCENARIO C

TN	8	
TD	2	
TW	1	
SL	E1-S1, E2-S1, E1-S2, E2-S2, E3-S3, E5-S6	cardinality=6
SO	SEQ, ANY	cardinality=2

Figure 23 – Event tree attributes for the Pickpocketing/Aggression scenario

	A-B	A-C	B-C
ΔTN	2	4	2
ΔSL	0	3	3
ΔTD	0	1	1
ΔSO	1	2	1
ΔTW	1	1	0
D	4	11	7

Table 4 – Differences among attributes of scenarios A, B and C

An overview of distances among attributes of event trees representing scenarios A, B and C is reported in Table 4.

As an example, in off-line operation, when inserting scenario B after A and C, the human operator sees the distances with scenarios A and C. In this case, he/she will be aware of the similarity with scenario A, since the distance is low (e.g. D_T could be set to 5) and could decide to keep only the original version (i.e. scenario A) since the variation would be automatically detected by the on-line heuristic engine based on distance.

In on-line operation, let us assume the ANY event of scenario A is detected. The expected behaviour will be as follows.

1. DETECT computes the attributes associated to the ANY composite event subtree (see below).

TN	8	
TD	2	
SL	E1-S1, E2-S1, E1-S2, E2-S2, E3-S3	cardinality=5
SO	ANY, OR	cardinality=2

2. DETECT computes the distances with all the (enabled and full) event trees in the Scenario Repository (see D row below).

	ANY-A	ANY-B	ANY-C
ΔTN	4	2	0
ΔSL	2	2	1
ΔTD	1	1	0
ΔSO	2	3	1
D	9	8	2

The computed distances correctly represent the recognised situation that, though formally belonging to scenario A, in absence of chemical warfare agent detection, is more similar to a situation of aggression/pickpocketing.

Given the possibility to get additional, but still appropriate warnings on possible forthcoming threats, the on-line operation is strategic to enrich the detection capabilities of the existing deterministic correlation engine. In particular, the described recognition technique addresses the imperfect threat modeling, due to human faults, as well as the possible missed detections, due to sensor faults.

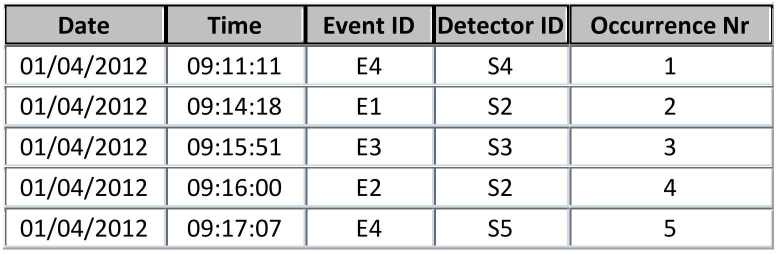
6.3 Real-time evaluation of alarm reliability

A practical application of the approach presented in Chapter 5 is the following. Let us consider the chemical attack scenario already addressed in the previous section (scenario A), which describes the drop of a CWA in a metro railway platform, represented by the event tree Figure 21. The scenario is built considering two intelligent cameras positioned at platform end walls, a microphone between them, two standoff detectors for CWAs positioned on the platform and on the escalator or concourse level. Let us assume to characterize the involved detectors with the FAR parameters reported in Table 5.

Detector ID	Detector Description	Event ID	Event Description	FAR
S1	Intelligent Camera	E1	Fall of person	0.25
		E2	Abnormal running	0.20
S2	Intelligent Camera	E1	Fall of person	0.25
		E2	Abnormal running	0.20
S3	Audio Sensor	E3	Scream	0.15
S4	IMS/SAW detector	E4	CWA detection	0.30
S5	IR detector	E4	CWA detection	0.27

Table 5 – FAR parameters of detectors used in chemical attack scenario

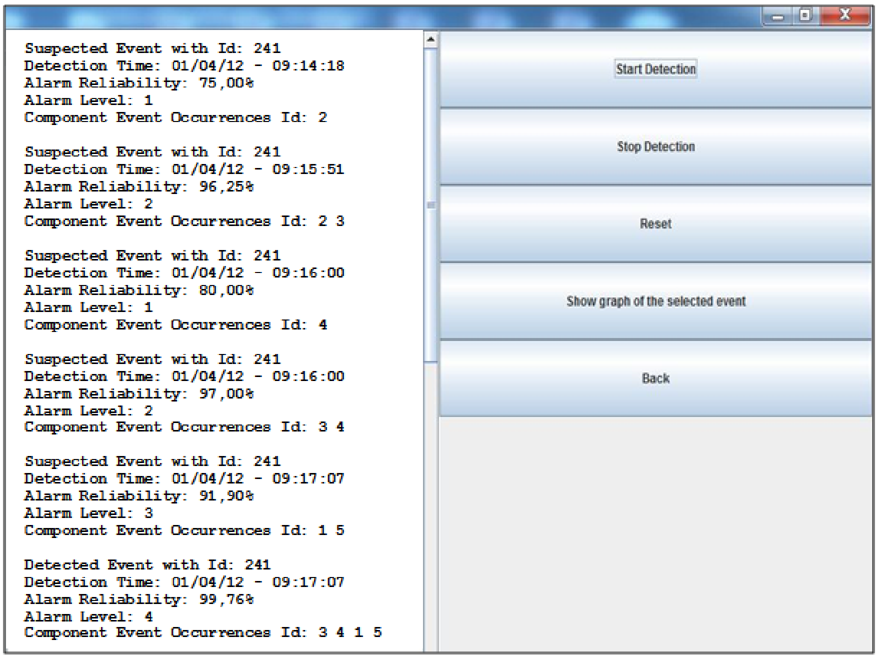
A possible set of basic event occurrences corresponding to a real CWA attack is listed in Table 6, which includes chronological aspects like the ones used in real PSIM log-files.



Date	Time	Event ID	Detector ID	Occurrence Nr
01/04/2012	09:11:11	E4	S4	1
01/04/2012	09:14:18	E1	S2	2
01/04/2012	09:15:51	E3	S3	3
01/04/2012	09:16:00	E2	S2	4
01/04/2012	09:17:07	E4	S5	5

Table 6 – A possible basic events chronology related to the CWA attack

When using DETECT to model the threat scenario (whose ID is assigned - for example - 241), with the Event Tree of Figure 21 and the parameters of Table 5, the output is reported in the screenshot in Figure 24: for each detected event, also the reliability level is reported, which is calculated in real-time using the BN approach. In the described example we have considered no uncertainty coming from the detection model (the confidence index of each operator used to build the event tree is set to its default value, i.e. 1).



```
Suspected Event with Id: 241
Detection Time: 01/04/12 - 09:14:18
Alarm Reliability: 75,00%
Alarm Level: 1
Component Event Occurrences Id: 2

Suspected Event with Id: 241
Detection Time: 01/04/12 - 09:15:51
Alarm Reliability: 96,25%
Alarm Level: 2
Component Event Occurrences Id: 2 3

Suspected Event with Id: 241
Detection Time: 01/04/12 - 09:16:00
Alarm Reliability: 80,00%
Alarm Level: 1
Component Event Occurrences Id: 4

Suspected Event with Id: 241
Detection Time: 01/04/12 - 09:16:00
Alarm Reliability: 97,00%
Alarm Level: 2
Component Event Occurrences Id: 3 4

Suspected Event with Id: 241
Detection Time: 01/04/12 - 09:17:07
Alarm Reliability: 91,90%
Alarm Level: 3
Component Event Occurrences Id: 1 5

Detected Event with Id: 241
Detection Time: 01/04/12 - 09:17:07
Alarm Reliability: 99,76%
Alarm Level: 4
Component Event Occurrences Id: 3 4 1 5
```

Figure 24 – Screenshot reporting alarms and their reliability values in real time

The real-time execution of the BN models also enables the possibility of using ‘dynamic’ FAR parameters, continuously updated using the feedback of the human operators in terms of confirmation of the alarms detected in the real on-the-field operation. In other words, for each event detected by a sensor, the statistical analysis of the ratio (‘false positive alarms’ / ’total number of alarms’), can lead to a proper update of FAR parameters and therefore to more reliable estimations with respect to the ‘static’ ones.

Chapter 7

Integration of DETECT with a PSIM system

7.1 Basic motivations

Physical security is ensured by monitoring and protecting users and the physical assets of a certain infrastructure (e.g. in a railway context, it includes stations, bridges, tunnels, ventilation shafts, depots, etc.). To that aim, PSIM systems (like the one described in [13]) are becoming a popular choice to integrate several heterogeneous sensing platforms. A typical PSIM system is distributed and its mission is to detect the relevant events occurring in the peripheral sites and eventually alert the operators in a Security Control Center, support the execution of the emergency procedures and/or activate automatic reactions. Since the PSIM system may generate a large amount of alarms which could overwhelm the personnel in charge of reacting to threats and suspicious events, in order to lower the false alarm rate and improve the detection reliability of threat scenarios, event correlation capabilities need to be integrated into the system. It is interesting to notice that, in several application domains (like the railway context), only a few intelligent surveillance systems effectively integrate and use event correlation. Some existing commercial products feature limited correlation capabilities, based on basic logic correlations, like in multi-technology detectors. Several reasons may motivate this matter of fact, including the need for light and efficient, and easy to use approaches in recognition of evolving situations based on a-priori knowledge of threat patterns, which in turn could be easily updated and integrated also by the security operators: that is an objective which is far from being trivial to achieve [32].

Therefore, the integration of DETECT with an existing PSIM system can represent an important step towards the development of a single cohesive platform able to:

- overcome limitations and doubts related to the PSIM system (as described in section 1.1.2 and 1.4);
- feature all the PSIM key capabilities (see Figure 2) and the added value they should provide, e.g. in terms of real-time data analysis and interpretation.

The advantage lies in the possibility to interact with a single interface, which comprise all the necessary information (coming from multiple sensing subsystems), including the ones related the occurrence of complex threats. The effectiveness of such an integration is twofold:

- reduction of the warnings (specially the unnecessary ones) sent to human operators, by means of a filtering on the detected events;
- prompt support in the activation of proper countermeasures.

Since each sensing subsystem, integrated in the PSIM platform (like the VMS), is not 100% reliable and the raised warning may not be significant or indicative of a real threat (and may not require a reaction), it is essential to report alarms only when strictly necessary or useful. The consequence is a simple, but effective support to the human operators in a control center. In fact, in the “standard” mode, each event/alarm detected by a sensing subsystem is reported to the operators (the possible aggregation of multiple events/alarms is thus performed at subsystem level and/or at junction box level). On the contrary, the PSIM system can be configured in the following way: standard mode only if the single event/alarm is critical and needs to be taken in charge by the operator immediately, otherwise the interface reports only aggregated alarms, i.e. the composite events recognized by DETECT. In turn, the latter can be filtered selecting the ones with “alarm level”, “alarm reliability” and “distance” above or below certain thresholds.

The additional advantage of such an optimization consists of reducing the workload of human operators, which can be more concentrate and quicker in taking in charge critical events, using the right priority. At the same time, the early warning of more complex events is strategic in the early adoption of more punctual countermeasures.

7.2 An integrated framework for railway protection

This section provides some details on the proposed integration between SMS (Security Management System) – developed by Ansaldo STS and used in railway transportation systems [13] – and DETECT. The objective is to enrich an existing PSIM system with basic, but effective reasoning capabilities.

SMS integrates intrusion detection, access control, intelligent video-surveillance and intelligent sound detection devices. The system is able to integrate also CBRNe (Chemical Biological Nuclear Radiological explosive) sensors to improve detection of terrorist attacks.

The SMS architecture (Figure 25) is distributed and hierarchical; a dedicated network provides reliable communication among the sites and an integrated management system collects the alarms and supports decision making. In case of emergencies, the procedural actions required to the operators involved are orchestrated by the SMS.

Data gathered from the heterogeneous sensing devices are processed by subsystems which generate the alarm events. Those alarms are first collected by peripheral control centers (Peripheral Security Places, PSP, e.g. positioned in the stations) and then centralized in a control center (Central Security Places, CSP, e.g. close to the traffic management center). Every security place (peripheral or central) can be provided with a SMS operator interface.

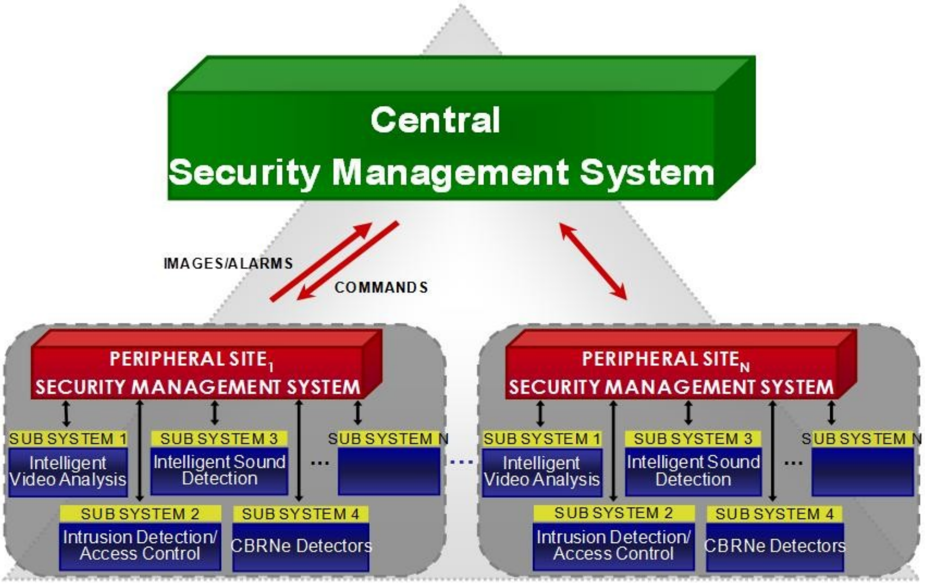

Central
Security Management System
IMAGES/ALARMS
COMMANDS
PERIPHERAL SITE$_1$
SECURITY MANAGEMENT SYSTEM
SUB SYSTEM 1
Intelligent Video Analysis
SUB SYSTEM 3
Intelligent Sound Detection
SUB SYSTEM N
SUB SYSTEM 2
Intrusion Detection/ Access Control
SUB SYSTEM 4
CBRNe Detectors
PERIPHERAL SITE$_N$
SECURITY MANAGEMENT SYSTEM
SUB SYSTEM 1
Intelligent Video Analysis
SUB SYSTEM 3
Intelligent Sound Detection
SUB SYSTEM N
SUB SYSTEM 2
Intrusion Detection/ Access Control
SUB SYSTEM 4
CBRNe Detectors

Figure 25 – SMS architecture

The events detected by the available sensorial subsystems are stored in appropriate repositories, both at the PSPs and CSP sites, and then collected in an Oracle DB (which corresponds to the Event History considered for DETECT).

In the integrated environment (see Figure 26), DETECT and SMS share the Event History database and communicate by exchanging warning messages (from DETECT to SMS) and possibly commands (from SMS to DETECT). The commands consist of specific feedback from human operators which can be used to refine or update the detection models handled by DETECT.

Therefore, the working logic is the following. On the one hand, SMS collects all alarms detected by heterogeneous sensorial subsystems and store them into the shared database. On the other hand, the engine of DETECT is fed by each new entry in the Event History. The interface mode with the database is asynchronous (i.e. by event-based queries, whenever a new event occurrence is stored in the Event History).

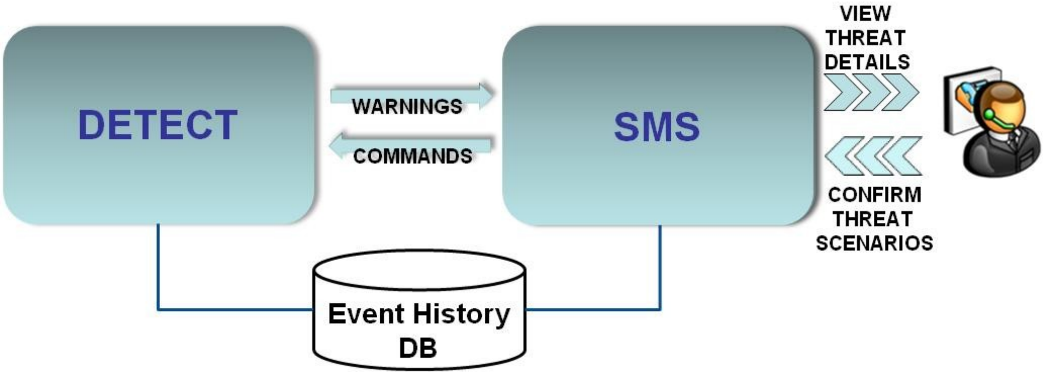


Figure 26 – Integration between DETECT and SMS by means of a shared database

The DETECT alert messages are then reported on the SMS operator interface, which include a dedicated view of all the detected alarms. Therefore, such a view gives to the security personnel information about the composite event that has been detected, i.e. semantic indication about the recognized situation (explosive in tunnel, chemical attack, etc.) and current phase according to the scenario evolution. However, such a view doesn’t include detections with low alarm reliability (if available), according to predefined threshold.

A possible issue to address is the following. If the detected scenario includes primitive events that have been already notified, SMS should drop them from the list of the alarms, after confirmation by the operator. The composite event can then be reported to the SMS interface. However the feature is not implemented in the current version.

Furthermore, depending on the specific configuration required, primitive events can continue to be shown in a hierarchical or tree structured view of the SMS interface. According to the parameters of the threat scenario (e.g. criticality level, detection reliability), the DETECT alarm may activate specific SMS procedures that will override procedures, possibly associated with primitive alarms. In fact, featuring an intrinsic lower level of reliability, alarms from single sensors need to be verified more carefully by the operator, while composite events could even trigger automatic countermeasures.

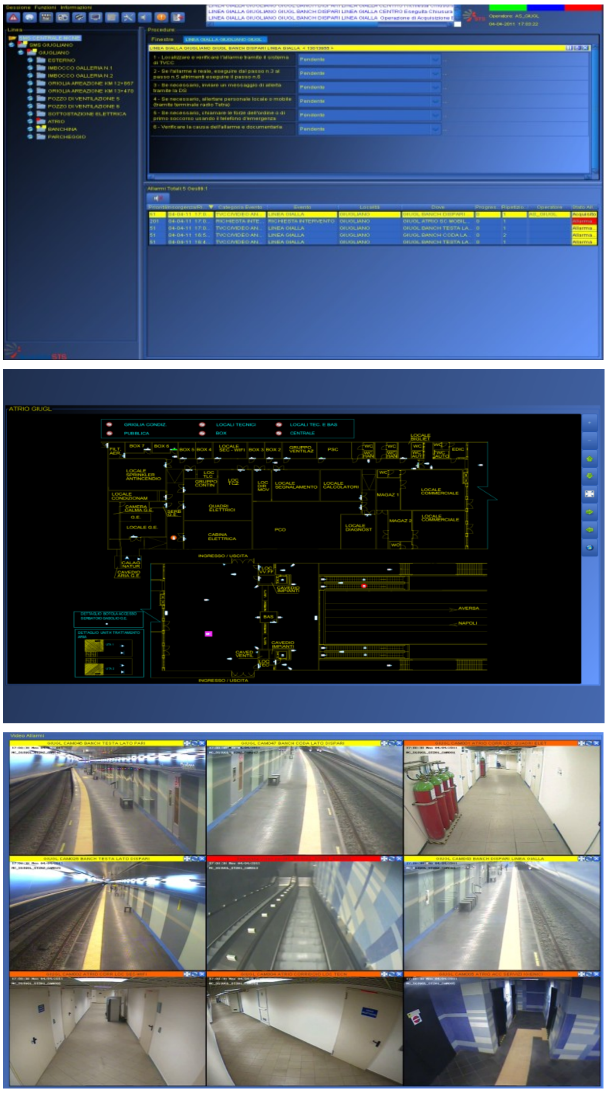
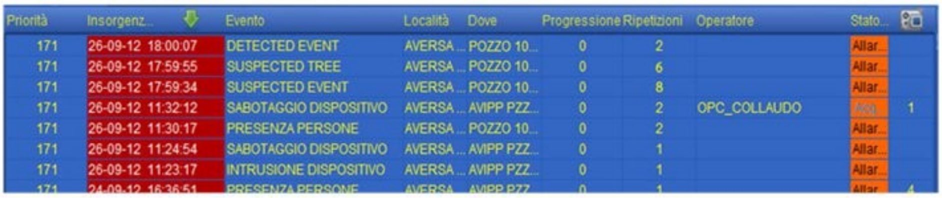


Figure 27 – An example of operator interface of the integrated system

Figure 27 shows an example of an operator interface including different screenshots of the integrated system. In particular, they show: the list of alarms with relative procedures (up), a vector graphics map which helps the operator to localize the source of the alarms (middle), the video streams automatically activated when an alarm is generated by smart-cameras or other sensors (down).



Priorità	Insorgenz...	Evento	Località	Dove	Progressione	Ripetizioni	Operatore	Stato...	
171	26-09-12 18:00:07	DETECTED EVENT	AVERSA ...	POZZO 10...	0	2		Allar...	
171	26-09-12 17:59:55	SUSPECTED TREE	AVERSA ...	POZZO 10...	0	6		Allar...	
171	26-09-12 17:59:34	SUSPECTED EVENT	AVERSA ...	POZZO 10...	0	8		Allar...	
171	26-09-12 11:32:12	SABOTAGGIO DISPOSITIVO	AVERSA ...	AVIPP PZZ...	0	2	OPC_COLLAUDO	Acq...	1
171	26-09-12 11:30:17	PRESENZA PERSONE	AVERSA ...	POZZO 10...	0	2		Allar...	
171	26-09-12 11:24:54	SABOTAGGIO DISPOSITIVO	AVERSA ...	AVIPP PZZ...	0	1		Allar...	
171	26-09-12 11:23:17	INTRUSIONE DISPOSITIVO	AVERSA ...	AVIPP PZZ...	0	1		Allar...	
171	24-09-12 16:36:51	PRESENZA PERSONE	AVERSA ...	AVIPP PZZ...	0	1		Allar...	4

Figure 28 – List of SMS alarms including DETECT warning messages

Figure 28 shows a detail of the list of alarms, which include the warning messages coming from DETECT. The first three rows of the list represent just a test performed to prove the correct acknowledgement of a:

- deterministic detection of a whole composite event (denoted with “DETECTED EVENT”);
- heuristic recognition of an event tree at low distance (denoted with “SUSPECTED TREE”);
- deterministic detection of a part of a composite event (denoted with “SUSPECTED EVENT”).

The proposed integration enables also an effective refinement process, based on the feedback of human operators (which confirm or not the alarms detected in the real on-the-field operation), stored in the shared database. The processing of this information (e.g. for each event detected by a sensor, the statistical analysis of the ratio “false positive alarms” / “total number of alarms”) can lead to a proper update of FAR parameters and hence to a continuous update of the detection models used by DETECT. Therefore, exploiting dynamic parameters, in line with on-the-field indications, instead of the static values, the overall correlation process is more reliable.

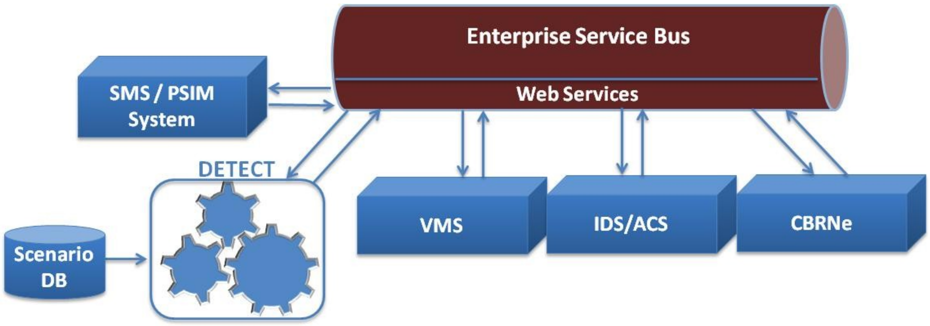


Figure 29 – ESB-based integration between DETECT and SMS

At the moment, the implementation of the integration is still in progress and we are moving towards a more advanced approach, consisting of a horizontal integration method. In fact, DETECT is configured for the data exchange with an Enterprise Service Bus (ESB), which implements the interaction and communication between mutually interacting software applications (like the ones of the single sensing subsystems) in a Service-Oriented Architecture (SOA). From this point of view, the ESB can collect on-the-field data and send them to DETECT via software, in such a way to feed its detection models (therefore, Event History is not a shared database, but an event queue handled by the ESB). Similarly, DETECT can send the recognized threat scenarios to the ESB, in order to display them into the interface of SMS / PSIM system. Such a solution provides more flexibility with regards to communication and interaction between applications, and bypass possible problems in the archiving of data in the shared Event History database.

7.3 Detection of distributed attacks

As stated previously, the SMS architecture is distributed and hierarchical. Thanks to the integration between DETECT and SMS, the same configuration can be repeated also for the DETECT architecture. The corresponding advantage makes possible the detection of simultaneous and distributed attack, which could not be recognized otherwise. In fact, only having a global view on the current status of all peripheral sites, it is possible to consider specific critical events. As a matter of fact, although they may be unlikely and/or apparently not meaningful from a local viewpoint, they may assume a different and concrete importance from a global viewpoint. This is especially true in case of a simultaneous occurrence of the same event in more places.

Assuming the simultaneous use of the correlation engine in each peripheral sites and in the main control center, it is possible to address strategic terrorist attacks (which often feature simultaneous strikes). In the CWA attack scenario considered in Chapter 6, if the control center detects its simultaneous (possibly partial) evolution in different subway platforms, then the evacuation of the involved stations and the block of train traffic could be triggered immediately. This approach can enable an advanced situation awareness, early warning and decision support. Accordingly it is possible to improve the impact of countermeasures in a significant way. This is the key to detect and respond to different, but increasingly widespread typology of threat scenarios, which are very difficult to recognize. The hierarchical architecture of the integrated monitoring system, including both SMS and DETECT, is functional for this purpose.

The implementation of such feature is quite simple with the available tools. In fact, on the one hand, the occurrence of a suspected threat scenario can be stored in the Event History (although it is not a primitive event occurrence), using the required format for the information encoding. On the other hand, the human operator can draw a new event tree (i.e. a sort of tree of trees), no longer fed by primitive events, but by composite events (or both of them).

Conclusions

The proposed solutions in this thesis are mainly focused on the management strategy of all the available resources (provided by the state-of-art) and on how to augment the capabilities of distributed surveillance systems in PSIM (with respect to the state-of-the-art), in order to better support security operators in responding to threats.

A general paradigm of augmented surveillance has been proposed to analyze all the components (which identify features on different levels) to be included, and to establish how to assemble them. Furthermore a model-based event correlation approach has been proposed to support situation recognition task in real-time and in different applications, thanks to a light, efficient and easy-to-use technique. Such features are crucial to assure the usability of the developed framework, i.e. DETECT, a detection engine mainly based on a specific event algebra to build Event Tree models. In fact, the latter satisfy the real-time solvability requirement and, at the same time, don't ask for a modeling expert to draw them, given their intuitiveness. Accordingly, this helps in making detection models easy to update and integrate by the security operators, reducing and simplifying the maintenance effort as well. The same considerations are not valid for more powerful formalisms (e.g. ANN, Petri Net), which are far from being straightforward to implement, control and update. However, the general architecture of the framework is suitable to accommodate different detection models, which could be used in parallel with the event trees.

Further contributions of this thesis are aimed at overcoming specific limitations of the introduced deterministic approach. To enhance the detection effectiveness of threat scenarios in PSIM, an event tree similarity analysis and a real-time distance computation have been proposed. The extension has several important advantages including non-exact tree matching, which enables heuristic detection both on-line, to achieve early warning and tolerance to imperfect modeling and missed sensor detections, and off-line, for a more rational use and management of the knowledge base provided by risk analysts.

To address the problem of uncertainty management in threat detection with PSIM systems, an additional feature has been introduced to evaluate the trustworthiness of the inferred alarms, which is always limited since modeling errors and sensor false alarms. In order to associate a reliability level to the detected threats, a real-time fuzzy correlation of sensor outputs has been performed by means of a Bayesian Network. Therefore, the formalism has been employed to complement the deterministic detection with probabilistic parameters, which characterize the sources of uncertainty (mainly the sensors and detection models themselves). According to the availability of these parameters, simple BNs enable a fuzzy logic whenever the output of the detection engine is not deterministic, but associated with a certain probability.

The decision making of the human operators is then supported by indications on suspected threat scenarios and related parameters (if available), like alarm level, detection reliability, and possible distances with other scenarios.

In addition to that, the integration of DETECT with a PSIM system, on the one hand, allows to achieve a superior situation awareness also on complex threats, which could be difficult to recognize; on the other, it allows to focus on fewer, but more significant event notifications, as well as to better and quickly discriminate between false and real alarms. Human operators can then orient their behavior accordingly, guided by customized event management procedures, possibly automated on the basis of the computed parameters.

We believe the main limitations of DETECT have been addressed and solved in a way that can be considered satisfactory to start using the framework in real surveillance applications. To that aim, we have already developed a set of threat scenarios relevant for metro railway contexts, some of which have been briefly addressed in this thesis for the case-study applications. We are presently finalizing the development of the integrated prototype, including DETECT and an existing PSIM system (i.e. the

Security Management System developed by Ansaldo STS), to demonstrate the effectiveness of the approach in metro railway environments.

Finally, we evaluating the possibility to define and implement ad-hoc techniques mainly aimed at the effective exploitation of PSIM operators' feedback, for the extraction of relevant information during the correlation process (as indicated by the augmented surveillance paradigm), e.g. in order to adjust models or create new ones. At the same time, the automatic 'learning' of uncertainty parameters, analyzing the number of confirmed alarms, is fundamental for a continuous update and fine-tuning of DETECT knowledge base, according to the real observations.

References

[1] Abidi, B. R., N. R. Aragam, Y. Yao, M. A. Abidi, *Survey and analysis of multimodal sensor planning and integration for wide area surveillance*. ACM Comput. Surv. 41, 1, Article 7, December 2008.

[2] Adams, W., Iyengar, G., Lin, C., Naphade, M., Neti, C., Nock, H., Smith, J., *Semantic indexing of multimedia content using visual, audio, and text cues*. EURASIP J. Appl. Signal Process, 2003(2): pp.170–185.

[3] Aghajan, H., Cavallaro, A., *Multi-Camera Networks: Concepts and Applications*. Elsevier, 2009.

[4] Akita, R. M., *User Based Data Fusion Approaches*. In Proceedings of the International conference of Information Fusion, Annapolis, Maryland, 2002.

[5] Alferes, J.J., Tagni, G.E., *Implementation of a Complex Event Engine for the Web*. In Proceedings of IEEE Services Computing Workshops (SCW 2006), Chicago, Illinois, USA, September 2006.

[6] Atrey, P. K., M. A. Hossain, A. El-Saddik, M. S. Kankanhalli, *Multimodal fusion for multimedia analysis: a survey*. Multimedia Syst, 16(6), 2010: pp.345-379.

[7] Atrey, P.K., Kankanhalli, M. S., Jain, R., *Information assimilation framework for event detection in multimedia surveillance systems*. Springer/ACM Multimedia Systems Journal, Vol. 12, No. 3, 2006: pp. 239-253.

[8] Bedworth, M., O'Brien, J., *The omnibus model: A new model for data fusion*. Aerospace and Electronics Systems Magazine, 15(4), 2000.

[9] Bisantz, A.M., Finger., R., Seong, Y., Llinas, J., *Human Performance and Data Fusion Based Decision Aids*. In Proceedings of the International conference of Information Fusion, Sunnyvale, California, 2009.

[10] Blasch, E., *Level 5 (User Refinement) issues supporting information fusion management*. In Proceedings of Information Fusion, Florence, Italy, 2006.

[11] Blasch, E., Plano, S., *JDL Level 5 Fusion Model "user refinements" Issues and Applications in Group Tracking*. SPIE Aerosense, 2002.

[12] Blasch, E., Plano, S., *Level 5: user refinements to aid the fusion process*. In Proceedings Multisensor, Multisource Information fusion: Archetetures, Algorithms, and Applications, Orlando, Florida, 2003.

[13] Bocchetti, G., Flammini, F., Pragliola, C., Pappalardo, A., *Dependable integrated surveillance systems for the physical security of metro railways*. In IEEE Procs. of the third ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC), 2009: pp. 1-7.

[14] Bossé, E., Guitouni, A., Valin, P., *An Essay to Characterise Information Fusion System*. In Proceedings of the International conference of Information Fusion, Florence, Italy, 2006.

[15] Buford, J., Jakobson, G., Lewis, L., *An approach to integrated cognitive fusion*. In Proceedings of 7th International Conference on Information Fusion, 2004.

[16] Candamo, J., Shreve, M., Goldgof, D. B., Sapper, D. B., Kasturi, R., *Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms*. In IEEE Transactions on Intelligent Transportation Systems, Vol. 11, No. 1, 2010: pp. 206-224. Available at: http://www.cse.usf.edu/~mshreve/publications/ITS.pdf

[17] Chakravarthy, S., Mishra, D., *Snoop: An expressive event specification language for active databases*. Data Knowl. Eng., Vol. 14, No. 1, 1994: pp. 1–26.

[18] Chakravarthy, S., Krishnaprasad, V., Anwar, E., Kim, S., *Composite Events for Active Databases: Semantics, Contexts and Detection*. In Proceedings of the 20th international Conference on Very Large Data Bases, September 1994.

[19] Ciardelli, L., Bixio, L., Ottonello, M., Cesena, M., Regazzoni, C. S., *Multi-sensor Cognitive Based approach to critical infrastructure protection*. In Proceedings of Third International Conference on Safety and Security Engineering (SAFE 2009), Rome, Italy, 1-3 July 2009.

[20] Cucchiara, R., *Multimedia surveillance systems*. In Proc. Of ACM Intl. Workshop on Video Surveillance and Sensor Networks, Singapore, November 2005.

[21] Cuppens, F., Miege, A., *Alert Correlation in a Cooperative Intrusion Detection Framework*. IEEE Symposium on Security and Privacy, 2002.

[22] Dasarathy, B. V., *Decision Fusion*. IEEE Computer Society Press, 1994.

[23] Dasarathy, B. V., *More the merrier. . . or is it?—sensor suite augmentation benefits assessment*. In Proceedings of the 3rd International Conference on Information Fusion (Fusion 2000), Vol. 2. IEEE, Paris, France, WEC3/20–WEC3/25, 2000.

[24] Dasarathy, B. V., *Information Fusion - what, where, why, when, and how?*. Information Fusion, 2(2), 2001: pp. 75-76.

[25] Davis, G.L., *CBRNE - Chemical Detection Equipment. eMedicine*, 2008. Available at: http://emedicine.medscape.com/article/833933-overview.

[26] Endsley, M.R., *Toward a theory of situation awareness in dynamic systems*. Human Factors 37(1), 1995: pp. 32–64.

[27] Endsley, M.R., *Theoretical underpinnings of situation awareness: a critical review*. M. R. Endsley, D. J. Garland (Eds.). Situation Awareness Analysis and Measurement. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2000.

[28] Esteban, J., Starr, A., Willetts, R., Hannah, P., Bryanston-Cross, P., *A review of data fusion models and architectures: towards engineering guidelines*. Neural Comput. Appl.14 (4), 2005: pp. 273–281.

[29] Fernandez, C., Baiget, P., Roca, F.X., Gonzalez, J., *Determining the best suited semantic events for cognitive surveillance*. Expert Systems with Applications, 38(4), 2011: pp. 4068–4079.

[30] Flammini F., Pappalardo A., Vittorini V., *Challenges and emerging paradigms for augmented surveillance.* In Effective Surveillance for Homeland Security: Balancing Technology and Social Issues, CRC Press / Taylor & Francis, 2013. To appear.

[31] Flammini, F., Mazzocca, N., Pappalardo, A., Pragliola, C., Vittorini, V., *Dealing with Uncertainty in Threat Detection with Event Trees*. Submitted to: Special Issue on Information Fusion for Safety and Security, Elsevier, 2013.

[32] Flammini, F., Pappalardo, A., Pragliola C., Vittorini, V., *A robust approach for on-line and off-line threat detection based on event tree similarity analysis*. In 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS'11), Klagenfurt, Austria, Aug. 30-Sept. 2, 2011: pp.414-419.

[33] Flammini, F., Mazzocca, N., Pappalardo, A., Pragliola, C., Vittorini, V., *Augmenting surveillance system capabilities by exploiting event correlation and distributed attack detection*. In Proc. 2011 Intl. Workshop on Security and Cognitive Informatics for Homeland Defence (SeCIHD'11), co-located with ARES'11, Springer LNCS 6908, 2011: pp. 191-204.

[34] Flammini, F., Gaglione, A., Ottello, F., Pappalardo, A., Pragliola, C., Tedesco, A., *Towards Wireless Sensor Networks for Railway Infrastructure Monitoring*. In: Proc. International Conference on Electrical Systems for Aircraft, Railway and Ship Propulsion (ESARS'10), Bologna, Italy, 19-21 October, 2010: pp.1-6.

[35] Flammini, F., Gaglione, A., Mazzocca, N., Pragliola, C., DETECT: a novel framework for the detection of attacks to critical infrastructures. In: Safety, Reliability and Risk Analysis: Theory, Methods and Applications, Martorell et al. (Eds), Procs of ESREL'08, 2008: pp: 105-112.

[36] Frost & Sullivan, Analysis of the Worldwide Physical Security Information Management Market, 2012. Available at: http://www.cnlsoftware.com/media/reports/Analysis_Worldwide_Physical_Security_Information_Management_Market.pdf

[37] Garcia, M. L., *The Design and Evaluation of Physical Protection Systems*. Butterworth-Heinemann, 2001.

[38] Gatziu, S., Dittrich, K.R, *Detecting Composite Events in Active Databases Using Petri Nets*. In Proceedings of the 4th International Workshop on Research Issues in data Engineering: Active Database Systems, 1994: pp. 2-9.

[39] Gehani, N. H, Jagadish, H. V., Shmueli, O., *Event specification in an active object-oriented database*. SIGMOD Rec., 21(2), 1992: pp. 81-90.

[40] Goldgof, D.B., Sapper, D., Candamo, J., Shreve, M., *Evaluation of Smart Video for Transit Event Detection*. Project #BD549-49, Final Report, 2009. Available at: http://www.nctr.usf.edu/pdf/77807.pdf (accessed March 6, 2012).

[41] Gouaillier, V., Fleurant, A., *Intelligent video surveillance: Promises and challenges*. Technological and Commercial Intelligence Report, March 2009.

[42] Hall, B. V., McMullen, S.A.H, *Mathematical techniques in multisensor data fusion*. Northwood, MA. Artech House, 2004.

[43] Hall, D., Hellar, B.D., McNeese, M., Llinas, J., *Assessing the JDL model: A survey and analysis of decision and cognitive process models and comparison with the JDL model*. In Proceedings of the National Symposium on Sensor Data Fusion (NSSDF), June 2007.

[44] Hall, D., Llinas, J., *Handbook of Multisensor Data Fusion*. USA: CSC Press LLC, 2001.

[45] Harris, C. J., Bailey, A., Dodd. *Multi-sensor data fusion in defence and aerospace*. The Aeronautical Journal 102 (1015), 1998: pp. 229–244.

[46] IPVideoMarket, PSIM Deployment Statistics. Report published on November 15, 2011. Available at: http://ipvideomarket.info.

[47] Jing-Ying, C., Liao, H.-H., Che, L.-G., *Localized Detection of Abandoned Luggage*. EURASIP Journal on Advances in Signal Processing, Article ID 675784, 2010.

[48] Kalgaonkar, L., Smaragdis, P., Raj, B., *Sensor and Data Systems, Audio-Assisted Cameras and Acoustic Doppler Sensors*. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2007.

[49] Kokar, M. M., Bedworth, M. D., Frankel, K. B., *A reference model for data fusion systems*. In Proceedings of SPIE Conference on Sensor Fusion: Architectures, Algorithms, and Applications, IV, 2000.

[50] Krausz, B., Bauckhage, C., *Automatic Detection of Dangerous Motion Behavior in Human Crowds*. In Proceedings of IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS 2011), 2011.

[51] Kumar, S., Meech, J.A., *A Hypermanual on Expert Systems*. Canada Center for Mineral and Energy Technology, 1994.

[52] Lewis, T.G., *Critical Infrastructure Protection in Homeland Security: Defending a Networked Nation*. John Wiley, New York, 2006.

[53] Mack, J.E., Long, G., *Physical Security Information Management*. Imperial Capital, May 2008.

[54] Makris, D., Ellis, T., Black, J., *Intelligent visual surveillance: Towards cognitive vision systems.* The Open Cybernetics and Systemics Journal (2), 2008: pp. 219–229.

[55] Martin, P.T., Feng, Y., Wang, X., *Detector Technology Evaluation*. MPC-03-154, Mountain-Plains Consortium (MPC), November 2003. Available at: http://www.mountain-plains.org/pubs/pdf/MPC03-154.pdf (accessed March 6, 2012).

[56] Mrad, A. Ben, Maalej, M.A., Delcroix, V., Piechowiak, S., Abid, M., *Fuzzy Evidence in Bayesian Network*. In Proc. Intl Conf. on Soft Computing and Pattern Recognition, 2011: pp. 486-491.

[57] Nilsson, M., *Human decision making and information fusion: extending the concept of decision support*. Technical Report HS-IKI-TR-07-002, Workshop on Information Technology, Halmstad University, 2007.

[58] Nilsson, M., Laere, J. V., Susi, T., Ziemke, T., *Information fusion in practice: A distributed cognition perspective on the active role of users*. In Information Fusion, Volume 13, Issue 1, ISSN 1566-2535, 10.1016/j.inffus.2011.01.005, 2012: pp. 60-78.

[59] Nilsson, M., Ziemke, T., *Rethinking Level 5:Distributed cognition and Information fusion*. In Proceedings of the International conference of Information Fusion, Florence, Italy, July 2006.

[60] Nilsson, M., Ziemke, T., *Information Fusion: A decision support Perspective*. In Proceedings of the 10th International Conference on Information Fusion, Québec, Canada, 9-12 July, 2007.

[61] Nilsson, M., *Mind the gap: Human decision making and information fusion*. Licentiate Thesis, Örebro University, 2008.

[62] Ntalampiras, S., *Audio Surveillance*. In Critical Infrastructure Security: Assessment, Prevention, Detection, Response, WIT Press, 2012: pp. 191-205.

[63] Ntalampiras, S., Potamitis, I., Fakotakis, N., *An Adaptive Framework for Acoustic Monitoring of Potential Hazards*. In EURASIP J. Audio, Speech and Music Processing, 2009.

[64] Ntalampiras, S., Potamitis, I., Fakotakis, N., *On acoustic surveillance of hazardous situations*. In Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP'09), Taiwan, Taipei, 19-24 April 2009.

[65] Pouget, F., Dacier, M., *Alert correlation: Review of the state of the art*. Technical Report EURECOM+1271, Institut Eurecom, France, Dec 2003.

[66] Prati, A., Vezzani, R., Benini, L., Farella, E., Zappi, P., *An Integrated MultiModal Sensor Network for Video Surveillance.* In Proceedings of the 3rd ACM International Workshop on Video Surveillance & Sensor Networks (VSSN05), Singapore, 6-11 November 2005.

[67] Qu, Y., Wang, T., Zhu, Z., *Remote Audio/Video Acquisition for Human Signature Detection*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2009.

[68] Roadnight, J., *Will Physical Security Information Management (PSIM) Systems change the Global Security World?*. CornerStone GRG Ltd Whitepaper, February 2011.

[69] Robins, D. B., *Complex Event Processing*. In CSEP 504, 2010.

[70] Rousseau, R., Breton, R., *The M-OODA Loop: A Model Incorporating Control Functions and Teamwork in the OODA Loop*. In Proceedings of the Command and Control Research symposium, 2004

[71] Sowa, J. F., *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co., Pacific Grove, CA, 2000.

[72] Spencer, B. F. Jr., Ruiz-Sandoval, M. E., Kurata, N., *Smart Sensing Technology: Opportunities and Challenges*. In Journal of Structural Control and Health Monitoring, Vol. 11, No. 4, 2004: pp. 349– 368.

[73] St. John, M., Risser, M. R., *Sustaining vigilance by activating a secondary task when inattention is detected*. In Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting, 2009.

[74] Steinberg, A. N., Bowman, C. L., *Revisions to the JDL Data Fusion Model*. D. Hall & J. Llinas (Eds.). Handbook of Multisensor Data Fusion, CRC press LLC, Florida, USA, 2001.

[75] Steinberg, A. N., *Foundations of situation and threat assessment*, in M. E. Liggins, D. L. Hall, J. Llinas (eds.), Handbook of Multisensor Data Fusion: theory and practice, second edition, CRC press, 2009: pp.437-501.

[76] Stotz, A., Nagi, R., Sudit, M., *Incremental graph matching for situation awareness*. In Proceedings of the 12th International Conference on Information Fusion (Fusion 2009), Seattle, WA, USA, 6-9 July 2009.

[77] Talantzis, F., Aristodemos, P., Lazaros, P. C., *Real Time Audio-Visual Person Tracking*. In Proc. of IEEE International Workshop on Multimedia Signal Processing, 2006.

[78] Tien, J. M., *Toward a decision informatics paradigm: a real-time, information-based approach to decision making*. IEEE Transactions on systems, man and cybernetics, Part C.33 (1), 2003.

[79] VidSys, *VidSys' Third Annual Security Survey – Research Brief*, October 2012. Available at:http://www.vidsys.com/resource/vidsys-third-annual-security-survey

[80] VidSys, *The parallels between PSIM for Physical Security and SIEM for Cyber Security*, Whitepaper, 2011. Available at: http://www.vidsys.com/resource/the-parallels-between-psim-for-physical-security-and-siem-for-cyber-security

[81] Vin, L. J., et al., *Information fusion for decision support in manufacturing: studies from the defense sector*. In International Journal of Advanced Manufacturing Technology, 35 (9-10), 2008: pp. 908-915.

[82] Wickens, C., Dixon, S., *The benefits of imperfect diagnostic automation: a synthesis of the literature*. In Theoretical Issues in Ergonomics Science, 8(3), 2007: pp. 201-212.

[83] Zhu, Z., Huang, T. S., *Multimodal Surveillance: Sensors, Algorithms and Systems*. Artech House Publisher, 2007.

[84] Zhu, Z., Li, W., Wolberg, G., *Integrating LDV Audio and IR Video for Remote Multimodal Surveillance*. In Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.3, 2005.